GENERALITY OF PERSONAL CONSTRUCTS

DISSERTATION

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

By

RACHEL JANE HAMILTON, B.A., A.M.
The Ohio State University
1952

Approved by:

OHIO STATE UNIVERSITY

Adviser
ACKNOWLEDGMENT

During both the preliminary and fixed-design stages of this research many people have given me assistance, some knowingly and some unknowingly. Throughout the entire period Professor George A. Kelly has continued to balance stimulating suggestions with a willingness to allow me to explore ideas and methods. The students in his research team, and the students and outside guests in his manuscript reading group have broadened my point of view by their discussions of personal construct theory and occasionally of the present research problem. Other professors on the faculty have helped through their readiness to discuss specific phases or problems of the research: Professors J.B. Rotter, J.E. Horrocks, H.B. English, J.R. Kinzer, H.E. Burtt, and R.J. Wherry of the Department of Psychology, and Professor D.R. Whitney of the Department of Mathematics. The students who served as subjects and those who served as judges have also given invaluable assistance. To all of these people I should like to express my deep appreciation.
TABLE OF CONTENTS

CHAPTER PAGE

List of Tables iv

I. Introduction 1

II. Historical Background 3

The Concept of Generality
Treatment of Generality in Various Theories
Treatment of Generality in Personal Construct Theory
Generality of Projective Test Data

III. Statement of the Problem 74

Origin of the Problem
Implications of Generality in the Psychology of Personal Constructs
Exploratory Research
The Problem as Revised
Experimental Hypotheses

IV. Testing Method and Procedure 104

Overview of Experimental Method
The Picture Test
The Repertory Test
Subjects
Testing Procedure
Pretest of Method and Procedure

V. Matching Method and Procedure 151

Alternative Treatments of Test Protocols
The Method of Correct Matching
Tests of the Null Form of the Hypotheses
Preparation of Materials for Matching
Judges
Training of Judges
Matching Procedure
VI. Statistical Analysis of Matching Data 191

Tests of the Null Hypotheses
Reliability or Comparability of the Judges
Differences between Pictures
Differences between Subjects

VII. Experimental Results 208

Reliability of the Judges
Tests of the Hypotheses
Characteristics of Matching in terms of Pictures
Characteristics of Matching in terms of Subjects

VIII. Discussion of Results 229

Some Sample Descriptions
Interpretation of Tests of the Hypotheses
Area of Permissible Application of Results
Observations about the Subjects
Observations about the Picture Test
Observations about the Repertory Test
Ways of Improving the Present Experiment
Relationship of the Present Experiment to Previous Studies
Implications of the Experimental Results for Personal Construct Theory and Its Use

IX. Summary and Conclusions 312

Bibliography 320

Appendix A - Supplementary Directions for Administering the Test 332

Appendix B - Instructions to the Judges 337
### LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLES</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Characteristics of Students Tested and Used As Subjects</td>
<td>135</td>
</tr>
<tr>
<td>II. Characteristics of Students Tested and Not Used As Subjects</td>
<td>136</td>
</tr>
<tr>
<td>III. Agreement of Matching by Two Judges</td>
<td>209</td>
</tr>
<tr>
<td>IV. Comparison of Two Judges' Accuracy of Matching Sets of Materials</td>
<td>210</td>
</tr>
<tr>
<td>V. Comparison of Patterns of Accuracy of Matching of Two Judges by Chi-square</td>
<td>211</td>
</tr>
<tr>
<td>VI. Evaluation of Reliability of Two Judges by Contingency Coefficient</td>
<td>213</td>
</tr>
<tr>
<td>VII. Evaluation of Accuracy of Matching Sets of Descriptions by Contingency Coefficient</td>
<td>216</td>
</tr>
<tr>
<td>VIII. Evaluation of Accuracy of Matching Sets of Descriptions by Critical Ratio</td>
<td>218</td>
</tr>
<tr>
<td>IX. Evaluation of Accuracy of Matching Sets of Descriptions by Chi-square</td>
<td>220</td>
</tr>
<tr>
<td>X. Comparison of Accuracy of Matching Various Pictures and Their Descriptions</td>
<td>224</td>
</tr>
<tr>
<td>XI. Characteristics of Matching in Terms of Subjects</td>
<td>226</td>
</tr>
<tr>
<td>XII. Product-Moment Correlations between Subjects' Scores on Various Hypotheses</td>
<td>228</td>
</tr>
<tr>
<td>XIII. Age Identification of Pictured Figures Described</td>
<td>270</td>
</tr>
<tr>
<td>XIV. Number and Type of Misperceptions</td>
<td>275</td>
</tr>
<tr>
<td>XV. Age Identifications in Terms of Repertory Test Titles</td>
<td>278</td>
</tr>
</tbody>
</table>
CHAPTER I

Introduction

The topic of investigation in this research may be said to have had two parents as well as many other ancestors, most of whom shall remain unnamed. One parent is theoretical psychology, particularly the psychology of personal constructs, which is now being developed by G. A. Kelly. The other parent is the observations made by the writer in an outpatient mental hygiene clinic. From both these sources have come questions about the generality, or range of applicability, of data obtained in sharply limited situations.

When psychological examinations are done in a clinical setting, it is nearly always assumed that the data obtained are not specific to the examination, but are applicable to situations outside the testing room. By some people the applicability is regarded as a fact not to be questioned. Others are aware of the assumption involved, and wonder about its validity. Of course no single statement can be made about the over-all applicability of test data to other situations as a whole. It is necessary to ask, instead, whether certain test data have some discernible relationship to other specific kinds of situations. In
our culture much of an individual's behavior both inside and outside the examination room is verbal behavior. One might ask, therefore, whether a client's verbalizations on a certain test or part of a certain test bear detectable similarities to his verbalizations at other times. If so, the psychologist is in a better position to anticipate the reactions of his client outside the clinic, to guess at probable reactions of other people to the client, and to note changes in the client during therapy.

As the name indicates, the psychology of personal constructs is a theory which emphasizes the advantages of knowing about and working with the conceptual systems of individual subjects, in contrast to dealing largely with so-called public or popular constructs. It is held that there is no precisely identical construct system held by all members of any cultural group, but rather that there are as many variations as there are individuals. Although two people use words in common, the constructs which their words symbolize are, to a certain extent, private or individual.

Within the psychology of personal constructs it is hypothesized that a person's non-verbal activities are related to his verbalizations. Eventually the hypothesis must be tested, but that is not the problem of the present...
investigation. It is believed that more precise predictions and controls over an individual's behavior can be made through use of his own construct system than through either knowledge of the situation alone or use of a construct system supposedly common to the majority of people from his cultural group.

Suppose we grant the personal nature of constructs. Now we may ask, "Do these personal constructs have sufficient breadth of application to make study of them worthwhile in the attempt to understand and predict individual behavior?" If most personal constructs are limited to a very few situations, then attempts of a psychologist to subsume the construct systems of his clients are likely to be so time-consuming as to be impractical. On the other hand, if certain personal constructs are used in reference to a large number of events, then study of the constructs should make a psychologist more efficient. Without a study of the breadth of application of personal constructs we do not know whether use of personal constructs elicited in a limited situation can lead to more accurate prediction than the use of public constructs.

Both the psychology of personal constructs and the writer's observations in a clinical setting lead to the question whether constructs used by a given individual in one situation are likely to be used by that individual in
other situations. Specifically, we may ask whether the constructs elicited by a "projective" technique are limited to the particular stimulus materials that elicited them or whether they have sufficient generality to appear in other situations. For purposes of the present research, "other situations" may be taken to refer both to other parts of the same test and to another test using a different technique. The two tests chosen for study here are a selection of pictures from the Murray Thematic Apperception Test and the Symonds Picture-Story Test plus the Kelly Repertory Test.

For practical reasons it was necessary to limit the investigation to study of a particular type of construct. Constructs about the "personality" or "character" of people were chosen for study because it is believed that they are important ones in daily interaction with others. Both the tests chosen for study here are capable of eliciting this type of construct. The Repertory Test was designed specifically to elicit constructs about people known to the subject. For the picture test new instructions were used so as to force the subjects to attend directly to the figures rather than to story themes. The tests and methods of administration will be described in more detail in chapters III and IV.
CHAPTER II

Historical Background

The Concept of Generality

The concept of generality is found in one form or another in many theoretical approaches. Here the term is used in its common English meaning to refer to "that which is general", i.e., "common to many; prevalent; extensive, though not universal" (128). As used in this paper the term generality will refer to the range of applicability of something. Although the concept is involved in many theories, from one theory to another there is variation in the idea of what is general. The discussion here will take up generality as it might be treated in examples of several types of theories. Generality will be discussed first as it might be formulated in learning theories. Then it will be examined in relation to psychoanalysis, to conceptualization theories, and finally in relation to neo-phenomenological approaches. Afterwards it will be presented in terms of personal construct theory.

Treatment of generality in various theories.

Although the word generality, as used here, and the term generalization, as used in several learning theories,
look similar, they do not refer to the same phenomena. The distinction will be seen more clearly in the course of the following discussion. Hull's stimulus-response theory of learning characterizes learning as the setting up of receptor-effector connections, which are strengthened through need-reduction. In addition to the primary receptor-effector connections which are reinforced, there are other such connections which are also strengthened.

"... The fact is that every reinforcement mediates connections between a very great number of receptor and effector processes in addition to those involved in the reinforcement process and represented in the conventional symbolism $H_R$. Several groups of such additional and indirectly established receptor-effector connections may be distinguished:

"1. The reaction involved in the original conditioning becomes connected with a considerable zone of stimuli other than, but adjacent to, the stimulus conventionally involved in the original conditioning; this is called stimulus generalization.

"2. The stimulus involved in the original conditioning becomes connected with a considerable zone of reactions other than, but related to, the reaction conventionally involved in the original reinforcement; this may be called response generalization.

"3. Stimuli not involved in the original reinforcement but lying in a zone related to it become connected with reactions not involved in the original reinforcement but lying in a zone related to it; this may be called stimulus-response generalization."

(46: 183)

It appears that when Hull wrote his book *Principles of Behavior*, he was envisioning the simultaneous reinforcement of a number of different pathways between receptor and effector organs. On the basis of Hull's statements two different guesses may be made as to what he would regard as
general or widespread. Let us take his phenomenon of stimulus generalization as an example. Although the term "stimulus generalization" is used, it is not actually the stimulus as such that is general. Using one set of constructs we may say that what is general is the reinforcement value of the stimulus. That is, a single stimulus reinforces a number of different connections. Using another set of constructs, we may guess Hull would say that what is general is an effector connection or motor response. That is, presumably one effector pathway is connected with a number of receptor pathways. Here one type of generality (reinforcement value) leads to another type of generality (effector connection).

Generalization is held to occur in terms of a gradient. In the case of stimulus generalization, a gradient is said to occur when a series of stimuli show decreasing degrees of generalization with the original or standard stimulus. Although Hull states that the characteristics of the afferent discharge are important (46:198), major attention is turned toward the characteristics of the stimulus as the determining factor in generalization (46:184). In Hull's system there is a continuous quantitative gradient of generalization based on similarity between stimuli in terms of physical attributes.

Although Wolpe accepts Hull's definition of the empiri-
cal phenomenon of generalization, their ideas of the process of generalization, and what is general, appear to differ somewhat. In several recent articles Wolpe has discussed stimulus-response learning theory in neurophysiological terms. He has stated that all sensory stimuli set up central neural excitation, and learning is correlated with the reduction of this excitation (132). He holds a view of quantitative gradients of generalization similar to that of Hull (133). He states that stimulus generalization occurs when two stimuli activate the same afferent neurones. In terms of this theory it is the afferent pathways that are general. In Hull's theory of stimulus generalization, on the other hand, generality presumably occurs in terms of common efferent pathways.

In opposition to Hull, Lashley and Wade have denied the existence of generalization (59). They say that in some experiments, in which human subjects were used, the appearance of stimulus generalization was given by the subjects' use of "habits of relational thinking". They state that the phenomena called stimulus generalization really represent a failure to distinguish between characteristics of the various stimuli and to associate them correctly with the conditioned response. The gradient, they hold, varies with degree of attention and "is created by or is a function of the organism and only secondarily, if at all, a property of the
physically definable character of the stimuli. Although they deny generalization, presumably they would not deny generality in the sense of occurrence of the same response to a number of stimuli. Apparently what is general in their idea is a response that may be either cognitive (in human beings and some animals, at least) or neuromuscular.

Still another view on generalization is presented by Razran (90). He holds that in conditioning experiments there are two kinds of generalization — "pseudo-generalization" (based on failure to discriminate between stimuli) and "true generalization" (which is "a positive capacity of the organism, an ability to generalize absolute characteristics of stimuli or objects"). Unlike Hull, Razran states that true generalization does not develop during the original conditioning but develops during the later testing for generalization. He believes that there are gradients of generalization but that they are crude, qualitative ones consisting of only a few steps. The gradients are formed when human beings and animals (at least the higher ones) "categorize or rate the new stimulus on some sort of crude similarity-dissimilarity scale". The categorizations may be changed, and they vary "much more with the organic dimensions of the organism than with the physical dimensions of the external stimuli in the CR situ-
What is general in this theory appears to be the response of categorization or rating which is interpolated between the various stimuli and the one motor response. This sounds much like interpolation of a cognitive response although not necessarily the kind of abstraction and generalization involved in concept formation.

Corresponding to Hull's phenomenon of generalization is Skinner's induction. Skinner states the Law of Induction: "A dynamic change in the strength of a reflex may be accompanied by a similar but not so extensive change in a related reflex, where the relation is due to the possession of common properties of stimulus or response." (110:32). Induction may be due to either similarity of stimuli (110:168) or proximity of stimuli (110:169) or both similarity and proximity. The term reflex, as Skinner uses it, refers not to neurology but to behavior, and specifically to the correlation between a class of stimuli and a class of responses. "An example of induction is the fatigue of a flexion reflex from one locus of stimulation through repeated elicitation of a reflex from another locus." (110:32) In this example it is the response of flexion which is general, i.e., which is elicited by stimulation at two different loci. Skinner's view is similar to Hull's in emphasis on generality in terms of a neuromuscular response or pattern of responses.
Generalization in conditioned response experiments
is regarded by Kluver (55) as a form of equivalence of stimu-
li or of responses. In the investigation of problem solv-
ing and transfer of training in monkeys Kluver used the
method of equivalent and non-equivalent stimuli, and in
1936 he suggested the use of that method in the study of
personality. He pointed out that the same stimulus may
elicit widely different responses and that the same response
may be given to widely different stimuli. Interest should
center, he said, not in the fact that there is equivalence,
but in the range of equivalence and the factors leading to
equivalence and non-equivalence. He pointed out that
"... certain responses are called forth by a very large
number of different objects, whereas in other instances the
number of objects is small." He holds that if stimuli
are found to be equivalent, we may conclude that they are
"similar" in some way, but it is similarity as perceived
by some organism. Both the properties of the external
world and the condition of the organism are important in
determining perceived similarity. He also points out that
in order to understand equivalence we must examine non-
equivalent situations:

"... we must look for something accounting for both
the equivalence and for the difference between the
groups of equivalent and non-equivalent stimuli. In
other words, there must be some property or constella-
tion of properties which is present in the group of
equivalent and not present in the group of non-equa-
Kluver's equivalence of stimuli and of responses may be considered the compliment of generality as used here. If a variety of stimuli elicit a certain response, the stimuli are considered equivalent, but it is the response which is general. And if a single stimulus can elicit a variety of responses, the responses are considered equivalent, but generality occurs in the ability of the stimulus to elicit responses over a wide range.

Mowrer's two-factor learning theory employs the terms generalization and induction, but in a sense different from Hull's and Skinner's uses of the terms. Mowrer uses "generalization" to mean "going from the particular to the universal" (77:335). He states his position as follows:

"... Through the two basic forms of learning - conditioning and problem solving - living organisms develop generalizations (or, if verbally stated, propositions) commonly known as attitudes and habits. An 'attitude' is a generalization that takes this form: 'A (the CS) has been followed by B (the UnCS) once, twice, three times, n times... A is always followed by B.' On the other hand, a 'habit' is a generalization that takes this form: 'This particular response has solved this particular problem once, twice, three times, n times... It will always solve this problem.'" (77:330)

Here, instead of solely neuromuscular responses, as in the theories described earlier, there may be cognitive responses. It is these cognitive reactions (attitudes and habits), that are thought to be general, i.e., are applied in a number of situations.
In contrast to the positions of the theories described above, there is no generality of responses in Thorndike's connectionism. There is generality only insofar as the same stimulus element appears in more than one situation. When transfer of training occurs, it is due to the presence of identical elements in the several situations and to the activation of specific stimulus-response bonds (43:19-51). Generality in a different form may be said to occur in his spread of effect. He holds that rewards intended to act upon one connection are found to act, also, on other connections (43:35). Here the effectiveness of the rewards is relatively wide, i.e., general.

A review of the probable treatment of generality within several learning theories leads to the conclusion that generality is held to occur in terms of reinforcement values (Hull, Skinner, Lashley and Wade), afferent neural pathways (Wolpe), cognitive reactions (Mowrer, Lashley and Wade, Razran), and stimulus objects or procedures (Thorndike). The area of generality postulated depends upon the rest of one's theoretical orientation. Now let us turn attention to the treatment of generality within psychoanalysis. Here we will deal first with Freudian theory and then with Adlerian psychology.

In Freudian psychoanalysis the concept of generality is handled through the principle of displacement. Hendricks
describes displacement as follows:

"Displacement is the representation in consciousness of a part or whole of the original unconscious phantasy by some associated substitute. It is, therefore, one of the main mechanisms by which an unconscious wish attains representation in consciousness, while the integrity of the unconscious, the preservation of the repression of the original idea, is maintained." (41:37)

Displacement may be of three types: one act may be substituted for another; one part of the body may be used to satisfy desires originally directed toward another part; and one thing or person may be substituted for another as the object of a person's love or hatred. "Transference" is merely a particular type of displacement, in which the psychotherapist displaces some other person (usually a parent) as the object of the patient's emotions. In this theory it is the desires or emotions of the person that are regarded as general, i.e., as capable of being satisfied by more than one act or object.

An early offshoot of Freudian theory, Adler's Individual Psychology, offers a new way of conceptualizing generality. According to Adler, individual acts cannot be understood apart from a person's whole pattern of behavior. This pattern, which he calls the "style of life" or "law of movement", is established during the preschool years, and is the way that a person tries to overcome his feelings of inferiority and strive toward a goal of superiority. (2:37-38). Generality comes into Adler's theory in the per-
vasiveness of this style of life, which determines all of a person's actions and capacities (2:15). The pervasiveness exists over both horizontal and vertical segments of a person's life. Except through treatment, people do not change their style of life, although their expression of it may change (3:3-14, 81). The unity of life is shown in Adler's statement that similar attitudes may be expressed and similar goals may be sought by two people through different actions, and the same actions in two people may express different attitudes and goals (2:82). Adler regards the style of life as an acquired thing rather than something innate, but he is vague about its systematic description. One might guess that he would say that there is generality of cognitive and perhaps emotional factors, but that there is no generality of stimulus-response relationships or neuromuscular reactions. Furthermore, the generality involves direction of responses by other responses, rather than merely reactivity of an organism.

Turning now to field theory, we may take Lewin's topological psychology (62, 63, 64) as an example. Generality may be considered in this theory from at least two directions. On the one hand generality occurs in terms of overall characteristics of the psychological field of any individual - for instance, atmosphere and amount of freedom -
as distinct from such specific items of the field as goals, stimuli, needs, and social relationships (63:793). Substitute activity and substitute satisfaction may be considered as another treatment of the question of generality. When a tension exists and a person cannot reach his original goal, substitute actions often occur (62:181). The occurrence of substitute actions is based partly on the characteristics of the activities, partly on the situation, and partly on the developmental state of the person (63:822). If the boundaries between need systems are weak, as in young children, and there is a great deal of communication between systems, then there can be a great deal of substitute satisfaction. We might say that generality occurs in terms of needs in the sense that each need can be satisfied by a relatively large number of behaviors or objects. This conception of substitute activities appears to be closely related to Freud's idea of displacement.

Another important group of theories stresses various conceptual and perceptual approaches. Here one group might be called particularization theories, that is, theories which emphasize precision of use of concepts. Another group might be termed level theories because of stress on levels of conceptualization. A third group includes the neo-phenomenologists with their stress upon the personal meaning
of experiences. In the particularization group Korzybski, Johnson, and Hayakawa have been prominent. Although these writers do emphasize levels of abstraction (and in this way show some resemblance to the so-called level theorists), their orientation in regard to the levels is in terms of the relationship between a given level and the precision or lack of precision with which a concept is used. Goldstein, Vigotsky, and Rapaport, who are grouped here as having "level" theories, are more concerned with the way in which conceptualization and the whole personality function. The neo-phenomenologists include such approaches as Lecky's self-consistency theory and the emphasis by men such as Snygg and Combs, and Raimy on the self-concept and related variables.

Korzybski, who may be taken as representative of the particularization group, points to the process of abstraction as fundamental. Every person, object, and event is different from every other one, he states, but human nervous systems are incapable of registering all the differences. Therefore, similarities "...are read into nature by our nervous system..." (57:166). The use of names and class terms necessitates the further ignoring of differences and makes for greater emphasis on similarities. Similarities are obtained only by ignoring differences, that is, by a process of abstracting. He attributes the difficulties of people
...We often live, feel happy or unhappy, by what actually amounts to a definition, and not by the empirical, individual facts less coloured by semantic factors. When Smith1 marries Smith2, they mostly do so by a kind of definition. They have certain notions as to what 'man', 'woman', and 'marriage' 'are' by definition. They actually go through the performance and find that the Smith and his wife, Smith2, have unexpected likes, dislikes, and particularities - in general, characteristic and semantic reactions not included in their definition of the terms 'man', 'woman', 'husband', 'wife', or 'marriage'. Characteristics left out in the definitions make their appearance. Disappointments accumulate, and a more or less unhappy life begins." (57:415)

He advocates remembering that "... the label is not the object, and that the object is not the event, etc." (57:416). All this leads one to conclude that Korzybski, if he spoke in terms of generality, would advocate less of it. He would say that the same words are applied to too wide a range of objects or situations and that as a result people try to act toward several objects or events as if they were identical, when actually they are quite different.

Goldstein’s work (29, 30) may be taken as representative of those theories in which level of conceptualization is primary. Whereas Korzybski points to abstraction as a process of ignoring differences and emphasizing similarities and would probably say that this leads to too much generality of terms and concepts, Goldstein regards abstraction as a pervasive "attitude" or capacity widely used by
Goldstein divides all thinking, language, and behavior into two sharply demarcated types, the so-called abstract and concrete attitudes. In the concrete attitude "... we are given over passively and bound to the immediate experience of unique objects or situations. Our thinking and acting are determined by the immediate claims made by the particular aspect of the object or situation..." (29:6). In the abstract attitude "...Our actions are determined not so much by the objects before us as by the way we think about them; the individual thing becomes a mere accidental example or representative of a 'category'." (29:6). The normal person can assume both attitudes according to the needs of the situation, but the brain-injured person is confined to the concrete attitude. The concrete attitude or capacity level, then, is regarded as extremely general, or widespread in the abnormal person. Although the normal person shows considerable generality of attitude, the degree is less than in the case of the abnormal person. Because of the extreme generality which they believe to be characteristic of the level of functioning, Goldstein et al (30) and Hanfmann and Kasanin (32) support the use of concept formation tests and prediction from a patient's performance on the tests to his level of thinking as a whole.

Neo-phenomenological theories are represented here by the views of Snygg and Combs, and of Lecky. Fundamental
to the view of Snygg and Combs is the "personal frame of reference", with an "attempt to observe behavior from the point of view of the individual himself" (113:10). As their basic postulate they hold that "All behavior, without exception, is completely determined by and pertinent to the phenomenal field of the behaving organism." (113:15) Behavior is held to be in a one-to-one relationship to the phenomenal field. The major factor in determining behavior is the need of the behaver, and the single basic need is "the preservation and enhancement of the phenomenal self" (113:58). The phenomenal self is defined as "...those aspects of the phenomenal field to which we refer when we say 'I'" (113:56). What a person does and how he does it are determined by the concept he has of himself. Although this concept can change, it is usually extremely stable because of the inertia of the organization of the phenomenal field and the selectivity of perception that is exercised by the phenomenal self. Goals and techniques for reaching the goals become differentiated when they aid in the satisfaction of the basic need. Once differentiated, these goals and techniques tend to persist, partly as a function of the degree of satisfaction of the basic need and partly as a function of the inertia of the field organization. Generality comes into the theory in the pervading influence of the concept of oneself on all behavior, and in the persistence
of certain goals and techniques in the service of the
basic need for self-preservation and self-enhancement.
Generality is extreme, for even behaviors which appear
inconsistent to an observer are stated to be determined
by the concept of self and the single basic need, if the
behaviors are examined from the point of view of the be-
haver.

Generality looms large also in the personality
theory of Lecky, who has stressed self-consistency.
For him stability over both short and long periods of
time exists in regard to goals.

"The fact is that, descriptively and mechani-
cally, there is no stability in behavior, and habits
do not exist. There is stability in respect to
goals or results, but none in respect to movement..." (60:7).

"It is our view that behavior is usually 'in
character' not because the separate acts are related
to one another, but because all the acts of an in-
dividual have the goal of maintaining the same
structure of values." (69:10)

The single purpose that Lecky conceives to be present in
all people is maintenance of one's own organization. Ac-
cording to him, a wide variety of behaviors may serve this
single goal. Here there is an extreme of generality in the
view of maintenance of self-consistency as all-pervasive.

Treatment of Generality in Personal Construct Theory.

Before a discussion of the use of generality in the
psychology of personal constructs (53), a brief descrip-
tion of relevant aspects of the theory is in order. The
basic assumption, Postulate A, is that "A person's psychological processes evolve toward what he construes to be an optimal anticipation of events." (53:29). This is an application to people as a whole of the use of prediction, so familiar in scientific thinking. It is held that each person forms predictions about the people, objects, and events around him in daily life. If the predictions are verified, he continues to operate under the concepts used in the predictions. If they are not verified, he may follow any one of several courses: for instance, he may change his prediction; he may continue to use the same prediction but regard the particular instance of failure of verification as an "exception"; he may decide that the particular instance was not an example of the concept used in the prediction but was an example of some other concept; or he may withhold judgment until he has tried out the prediction in other situations.

The predictions are held to be made in terms of constructs. In this system a construct is defined as follows:

"A construct is a way in which some things are alike and yet different from others. In its minimum context a construct would be a way in which two things are alike and different from a third. It should be kept in mind that the way in which the two things are like each other should be the same as the way in which they are different from the third" (53:75).

Construing a given event means hypothesizing that the event can be included in a certain class of events made up of at
least one other event to which it is similar in some re-
spect and at least one other event from which it differs
in the same respect. A construct is conceived to have two
endpoints (53:87). In many cases the construct is de-
scribed by a word and its antonym. Sometimes a person
can verbalize only one end of a construct; then instead
of giving an antonym he will say "not _______." (Here
the blank space stands for the word used to symbolize the
contrasting end of the construct.) However, a construct is
not necessarily verbalizable at all (53:32). A person may
be able only to describe or point to specific examples sub-
sumed by the construct. When a person does have a label for
a construct, that label is regarded as only an element in
the construct, not as the construct itself (53:119). When
we communicate with other people, usually we deal only in
terms of labels, and that fact may cause confusion. The
writer further views constructs as dimensional in nature
and in this view she probably puts more emphasis on a scale
than does Kelly. Kelly sees constructs as dimensional, but
currently holds that the dimensions acquire their meanings
basically as dichotomies; he believes that the median values
of a scale are likely to be loose or ambiguous. The writer's
stand is based on administrations of the Repertory Test
(to be described in Chapters 3 and 4) during both the pre-
iminary and final stages of this research; frequently sub-
jects spontaneously placed acquaintances along scales when asked to categorize them in terms of their own constructs. Constructs are held to be organized into construction systems. "A construction system is an arrangement of constructs in which the common features of certain constructs are, in turn, recognized as separating them from other constructs within the system. A construction system involves constructs of constructs and constructs of constructs of constructs." (53:32)

The terms construct and concept have been used in many other ways different from the use of the term in PPC. In order to clarify the matter, some of the other uses will be reviewed and compared with the present usage. In a recent article (126) Vinacke concludes that none of the current definitions is entirely satisfactory. He summarized the essential features which most of the uses of the two terms have in common:

"1. Concepts are not direct sensory data but something resulting from the elaboration, combination, etc., thereof....

"2. ...concepts depend upon the previous experience of the organism.

"3. Concepts are responses which tie together, or link, or combine discrete sensory experiences. This condition may be demonstrated by showing that an individual responds to different stimuli in the same way.

"4. It may be inferred that such ties or links are symbolic in nature; that is, the same response stands for a variety of data. In the human organism
this response is usually a word, and the word ties together different experiences with the same object, experiences with different objects somehow related to each other, the emotional responses aroused in these experiences, etc.

"5. On the side of the internal processes of the organism, concepts represent selective factors."

Because he believes that none of the definitions is satisfactory, Vinacke does not report any. In view of that fact, several of the uses will be summarized here. The statements below will be limited to little more than brief answers that various writers have given to the question, "What is a concept or construct?"

1. One of the oldest points of view is found in Aristotelian thinking. Here a concept would probably be called a class or category of objects or events. Membership of an object in a class was held to define its essential nature, and therefore its behavior, in all respects. Classes were often thought of as dichotomies with no gradation between them (62:4). An approach to this use of the term concept in recent times is shown by Thorndike; he appears to think of concepts as absolute categories, for he states that "Good concepts are ones that enable a person to classify things or events without mistakes." (121:143). However, he does not go so far as to imply that membership in a class determines the behavior of an object in all respects.
2. Probably the most common usage in recent times is the idea that a concept is an abstraction of similar aspects of two or more objects. The term was used in this way by Hull (47), Heidbreder (39), Rapaport (89), and others. Some writers, e.g., Rapaport, specify that a given word (symbol) may have varying content for different people. This is in marked contrast to the Aristotelian stand that the characteristics of any object are determined by its class membership.

3. A smaller number of writers, according to Woodworth (134) and Vinacke (126) have treated concepts as predictions of similarities in several objects. Smoke (112) follows this usage. Vinacke states that this has hardly ever been used as the sole explanation of concept formation, and that an aspect of this use - active participation by the subject in the formation of the concept - is of more importance.

4. In recent years such men as Stevens (116), Marx (71), and Spence (114) have characterized operational definition of concepts as essential. Stevens has quoted Bridgman as stating in *The Logic of Modern Physics* "... in general, we mean by any concept nothing more than a set of operations; the concept is synonymous with the corresponding set of operations." Walters and Pennington (71:32) have criticized this stand, saying that any change in operations would mean a new concept, and therefore a multiplicity of concepts would result. In rebuttal Stevens states that "criteria defining the con-
ditions for class-inclusion" are the operational tests required. If we accept definition of criteria of classes as operational tests, then this usage of the term concept becomes very similar to the earlier stated abstraction of common elements in several objects or events. The common elements may be considered to be the criteria for inclusion in a concept.

5. Sometimes the term construct is used to refer to an abstraction of the empirical relationship between stimulus and response variables. This type of construct MacCorquodale and Meehl (66) have called an intervening variable. Examples of this type of construct are Tolman's "demand", Skinner's "reserve", Hull's "habit strength", and Lewin's "valence".

6. Sometimes the term construct is used to refer to a hypothetical process or entity intervening between observed stimulus and response variables. MacCorquodale and Meehl (66) use the phrase "hypothetical construct" for this type, but Spence (115) terms this type an "intervening variable". This kind of construct involves the addition of something not observed and should be consistent, as MacCorquodale and Meehl point out, with whatever other relevant
knowledge exists. Examples of hypothetical constructs are Murray's regnancies, Allport's biophysical traits, and Freud's libido.

The question of the meaning of signs, which Osgood discusses in a recent article (82), is closely related to the matter of concepts. The writer is not sure whether Osgood would regard these two problems as the same. Osgood bases his "mediation hypothesis" on Hullian theory. He states that the meaning of a sign is the composition of the mediation process, i.e., a fraction of the total object-elicited behavior which finally is elicited by the sign alone after the sign has been present as a part of a total stimulus complex. To change the meaning of signs, behavior in relation to the objects signified must be changed. The meanings which various people have for the same signs differ in accordance with their behaviors toward the objects represented by the signs.

The term "construct", as used in personal construct theory, resembles most closely use number 3, above. Both emphasize the importance of abstraction of similar aspects of two or more objects or events and, also, both stress that concepts arise as predictions. However, there are certain differences between the two usages. In personal construct theory, stress is on not only the origin of con-
structs as predictions but also their continued use as predictions. Furthermore, the predictions are held to be concerned with differences between objects as well as similarities between them. If we accept Stevens' statement that the criteria defining the conditions for class-inclusion are operational tests, then the use of the term construct in PPC may be considered to be a particular example of an operational definition. The present use differs from both the fifth and sixth uses, above, in that the latter involve abstractions of the relationship between stimulus and response variables, while the former involves relationships between several response variables. The present view also is opposed to the seventh usage. In terms of Osgood's theory, meanings are in accordance with behavior, while in terms of PPC behavior is in accordance with meanings.

In PPC constructs are held to arise through anticipations or predictions of events.** Objects and events about which similar and contrasting predictions are made - particularly if the predictions are validated - are grouped together as elements in a construct. In the beginning, the abbreviation PPC will be used to signify the phrase psychology of personal constructs.

** No extended discussion of the development of constructs will be given here, since that is not the primary problem of this research.
the construct may or may not be named verbally. Sometimes, particularly in the early stages of the development of a construct, one of the elements may be used as the name of the whole construct. For example, a person may describe a woman as "motherly" and mean that that woman's behavior is similar to the behavior of his own mother or other mothers, while, at the same time, her behavior contrasts with that of at least one other person. Sometimes, particularly in the later development of a construct, some other word that was not originally the name of one of the elements may be used to stand for the entire construct. Thus a person might use the term "nurturing" instead of "motherly".

The principle of revision of constructs through experience is stated in Corollary A-1 (52:34): "A person's construct system varies as he successively construes experience." As the person's predictions are invalidated, his constructs may change. He may, for instance, come to regard some behavior other than nurturance as motherly, or he may decide that some person who was once an element in the construct does not fulfill the criteria of a motherly person and should no longer be included as an element in that construct. If a person uses any given term both today and tomorrow, we should not assume that he means exactly the same thing on the two occasions. And if he uses
the same term today and a year from today, still less should we assume that his meaning is the same, since the longer time interval gives an opportunity for more experience capable of leading to a change in the construct. However, in practice it is often necessary to proceed as if the other person is using the same construct when he uses a word on two different occasions.

Since constructs are held to be formed through experience and since the experiences of people differ, it is reasonable to assume that the constructs of two people differ in some measure, even though they use the same label for their constructs. Even if two people participate in or witness the same event, they will probably attend to somewhat different aspects of the event (52:5). And most certainly each will view another person as an observer or participant in the event. Therefore, their constructions, of the event may be expected to vary somewhat. (53:40-41). Corollary A-2 summarizes this view in the statement that "Persons differ in their construction of experience" (53:40). The variation can be demonstrated easily either by asking a number of people to define certain words or by asking them to state what they consider to be the antonyms of certain words. In the case of many words variation will be found in their definitions and their antonyms. This has been demonstrated, for instance,
by Grayson & Tolman in a study of the concepts of clinical psychologists and psychiatrists (31). Despite the differences in experience, at least in any one sub-culture, various people do tend to have some experiences in common. Thus the constructs used by different people are not completely unrelated to each other. Popular thinking tends to stress the similarities in the use of terms by various people. PPG recognizes that there are similarities, but emphasizes the individual or personal nature of constructs.

It is well to recognize variations in constructs both between persons and within any given person from time to time. However, variation is not the only factor to be considered. A certain amount of consistency is to be expected, both intra-individually and inter-individually. Inter-individual consistency has already been mentioned briefly. More attention will be given in this research to intra-individual consistency.

By consistency of constructs we mean compatibility of predictions. In defining the term consistency, Kelly has stated that:

"When we hold to two views which are consistent with each other we expect to choose similar, or at least compatible, courses of action under them. The two views are inconsistent if they require us to perform the impossible feat of riding off in opposite directions at the same time. They are inconsistent if they lead us to anticipate two incompatible events."

Intra-individual consistency of constructs is based on the
repetition of somewhat similar events and on the fact that a person construes new experiences in terms of constructs formed in relation to previous experiences. Unless a single new experience has unusual vividness, it may be expected to lead to a little change only in the person's construction system.

The relative stability of a person's construction system is expressed in Postulate B: "Variation in a person's construction system is subordinate to certain more permeable aspects of his system." (53:60) Permeability is defined as follows:

"A construct or an aspect of one's construction system can be called permeable if it is so constituted that new experience and new events can be discriminately added to those which it already embraces... It is under theregnancy of such constructs that subordinate aspects of one's construction system can be systematically varied without having his whole psychological house fall down on him." (53:61)

This postulate places limits to the variability of a person's construction system. The limits are the permeability of his constructs. It is expected that the construction systems of most people will be flexible enough to embrace the majority of new experiences which they have, and that changes of constructs will be made within the framework of the existing system rather than by the wholesale rejection of the current system and adoption of an entirely new system.
This is not to say that no one ever uses, in succession, contradictory constructs. Indeed, contradictory sub-systems may be used (Corollary B-1)(53:62), but their use is limited by superordinate systems (Corollary B-2) (53:69). Although parts of a person's total construction system may appear contradictory when looked at in relative isolation, they may be seen as consistent when viewed as aspects of a larger system, i.e., a system at a higher level of abstraction.

Suppose, now, we compare intra-individual with inter-individual consistency. It seems reasonable to assume that on the average, the experiences of a single individual will have more in common than the experiences of different individuals. (It should be remembered that experience is regarded as a construed event, not just an event.) In the case of any one person, events are looked at through the eyes of a single individual against a single background of previously construed events. In the case of many individuals, events are looked at against many backgrounds that may be quite diverse. Now suppose that a number of people are asked to construe the same set of events. It is to be expected that, although the construings of the various people may show some features in common, their construals will also differ from each other; but on the other hand there will be some commonalities in the way any one
individual construes even different events. Thus if a psychologist familiar with personal construct theory is given records of the constructions of several events by a number of individuals, he might be expected to be able to distinguish the several constructions of each individual with a better than chance degree of accuracy. This is essentially the question to be examined in the present research.

The discussion of personal construct theory up to this point has been intended to serve as a foundation for understanding the treatment of generality within the framework of the psychology of personal constructs. The ideas about generality which will be presented here will be those of the writer, but an attempt will be made to relate them to the current thinking of Kelly in his development of PPC.

The phrase generality of constructs refers here to the extensiveness of use of constructs. In terms of PPC the generality of a construct is expressed by all the elements subsumed by (included within) that construct by a specific person at a specific time. The elements at both ends of the dimension are included within the construct. For example, if a person applies to other people a dimension in which "understanding" describes one end and "rejecting" describes the other end, people conceptualized as either
understanding or rejecting are included as elements within the construct. The generality of the construct is such as to include all the people described as either understanding or rejecting, but not to include people not described as either understanding or rejecting.

The generality of a construct refers to the extensiveness of its use by a given individual. It does not refer to the extent of use of a dimension by a number of people. That is, the orientation is idiographic rather than nomothetic. If one wished to determine the generality of a construct, the task would involve obtaining a record, through spontaneous conversation, guided interview, and tests, of all the people, objects, or events to which a given person applied a particular construct at a particular time. The person using the construct would have to be the judge as to whether specific events were included in the construct, although the interviewer or examiner might use questioning to encourage consideration of possible elements by the other person.

The phrase "degree of generality" refers to the number of elements included within a specific construct by a specific person at a specific time. The more elements are included within a construct, the more general that construct is said to be. The less widely a construct is used, i.e. the smaller the number of elements in a construct,
the less general or more specific the construct is said to be. Since a construct, by definition, subsumes at least three elements, three is the smallest degree of generality that a construct can have. It is impossible for a construct to have zero degree generality.

In reference to a single construct no scale of generality is possible. A scale of generality is applicable only when several constructs are compared. The construct either is applied to a given element or it is not applied to that element; there is no partial inclusion of the element within the construct. A person may place a given element along a scale between the two end-points of the construct in order to express the element's greater resemblance to one end of the dimension than to the other, but this is not a scale of generality. In one sense generality may be scaled if one compares the number of elements included by a specific person in each of two or more different constructs. The constructs may be listed in order of the number of elements subsumed, and the construct which subsumes the largest number of elements may be said to have the greatest degree of generality.

Generality of constructs should be compared with and differentiated from certain other variables in the psychology of personal constructs. The concepts of generality and consistency are related, although they are not synonym-
mous. Kelly has stated that "When we hold two views which are consistent with each other, we expect to choose similar, or at least compatible, courses of action under them." (52:66) Likewise, when the generality of a construct is such as to include certain elements, we should expect the person holding the construct to behave in compatible fashion toward those two elements, insofar as his behavior is governed by that construct rather than by other constructs of which the elements may be part. The logic, however, can be reversed only in part. It cannot always be said that if a person's actions in regard to two events are consistent, then both events are elements subsumed by the same construct. For instance, a patient in a clinic may shun people whom he sees as "clinging" as well as those whom he calls "authoritarian". The lack of subsumption of the two elements by the same construct is apparent if we remain on a relatively low level in the hierarchy of subordinate-superordinate constructs. But if we look for constructs on a higher level of superordinality, then in many instances, and perhaps in all instances, we will find that both events are elements in the same construct. In the particular example cited, both "clinging" and "authoritarian" might be subordinate constructs subsumed by the construct "demanding person". Our hypothetical person might also shun boa constrictors, subsuming both boas and demand-
ing people under the construct "harmful" at a still higher level of superordinality.

The phrases "consistency in the use of constructs" and "generality of constructs" are not identical in meaning, although they may at first appear to be so. Consistency in the use of constructs means that at various times the construct, as used by any given person, "means the same thing", or is composed of the same or closely similar elements. Generality of constructs refers, theoretically at least, to the subsumption by a construct of a number of elements at a single instant or over a very brief period of time. The relative distinction between cross-sectional and longitudinal segments of time is here brought to the fore. In actual practice it is often necessary to proceed on the assumption that generality and consistency in use are the same. When a person uses a verbal symbol on two different occasions, we must usually (unless we have evidence to the contrary) act as if the construct is the same and the speaker is merely using the label in regard to two of the elements subsumed by a single construct. In other words, we must act as if there is longitudinal generality of constructs.

The relationship between the subordinate-superordinate hierarchy of constructs and generality of constructs has been hinted at. Any superordinate construct is considered
to be more general than the constructs which it sub-
sumes. Thus "furniture" is considered to be more gener-
al than "table" or "chair". Although any superordinate
construct is considered more general than its own sub-
ordinate constructs, it is not necessarily more general
than any subordinate construct. Some subordinate con-
structs may be more general than some superordinate con-
structs. Superordination-subordination are not absolute
categories into which constructs may be divided. Although
in many cases it is safe to assume that a given construct
is superordinate or subordinate to another construct be-
cause of common experiences in a certain culture, often
such an unchecked assumption will lead to misunderstand-
ing. It should also be remembered that subordination and
superordination are relative: a given construct may be
both subordinate to one construct and superordinate to
other constructs. It is for these reasons that the re-
lationship between generality and the subordinate-super-
ordinate hierarchy can be stated only in terms of constructs
within a single hierarchy and not in terms of constructs
which are parts of different hierarchies.

The statement was made earlier that variation in a
person's construct system is subordinate to more permeable
aspects of the system. At that time it was stated that
"A construct or an aspect of one's construct system can be
called permeable if it is so constituted that new experience and new events can be discriminatively added to those which it already embraces. (53:61) Both permeable constructs and general constructs subsume a number of elements, but constructs may have a certain degree of generality while being either permeable or impermeable. Permeability refers to whether or not a construct is "open" or "closed" to the addition of new elements; generality refers to the number of elements already included in the construct at any one time. A permeable construct may become more general by the addition of new elements, while an impermeable construct may retain a constant degree of generality or may, perhaps, show a decrease in degree of generality through reduction in the number of elements in the construct. The latter may occur when, for example, a person changes his ways of looking at other people during therapy.

Kelly distinguishes three types of constructs: preemptive, constellatory, and propositional constructs. These are defined as follows:

1. A **preemptive construct** is one which preempts its elements for membership in its own realm exclusively. "Anything which is a ball can be nothing but a ball." (e.g., species names)

2. A **constellated construct** is one which fixes the real memberships of its elements. (Anything which is a ball has also got to be... ...) (e.g., stereotypes)
3. A propositional construct is one which does not disturb the other realm relationships of its elements. 

"Any roundish mass may be considered, among other things, as a ball." (53:107b)

One cannot tell from a verbal symbol alone whether, for a particular individual, a construct is preemptive, constellated, or propositional. His use of the symbol is what is important. Representatives of all three types of construct may have the same degree of generality if they subsume the same number of elements. That is, the type of construct has no necessary relationship to the degree of generality. In practice it may be expected that constellated constructs will usually be more general than preemptive constructs, and propositional constructs will usually be more general than constellated constructs.

There are close similarities, but also some differences between the concept of generality as used here and the concept of comprehensiveness as used by Kelly. According to his usage

"Comprehensive constructs are those whose likeness end subsumes a relatively wide variety of events. They are not necessarily highly regnant or superordinate constructs, for the events which they subsume may all be relatively low on the superordinate-subordinate scale. A permeable construct tends to move in the direction of comprehensiveness because its open-endedness enables it to embrace more and more elements in its context as time goes on. A comprehensive construct is likely to be one which has been in use a long time, although, in certain cases, some manics for example, there is a dilation which
sometimes appears to bring forth a matrix of comprehensive constructs in a relatively short time....

"Actually, as we see it, a comprehensive construct is one which cuts across many other construct lines. The 'variety' in the elements is established by the person's having otherwise distinguished them as being different from each other by means of other constructs. Thus when we use the term 'variety', we are referring to a 'variety' within the person's own construct system. Thus a constellatory construct would tend to be less comprehensive than a propositional construct which embraced precisely the same elements. The constellatory construct tends to fix its elements with respect to other realm memberships and hence they cannot be construed in the same variety as they would otherwise. A wholly preemptive construct could, of course, not be comprehensive at all." (53:599-600)

The relationship of comprehensiveness and generality to permeability and to superordination-subordination are conceived to be similar. The two concepts differ chiefly in terms of their relationship to number and 'variety' of elements in the constructs, and so, also, in relation to the three types of constructs. If a construct subsumes a large number of elements, it is considered to have a high degree of generality even if a person describes all the elements at one end of the construct dimension as identical in all respects. For instance, if a person takes into consideration all the hairs in the world and considers all of them to be identical, then for that person, "hair" is a construct with a high degree of generality, even though the construct hair is not considered to be a comprehensive one (53:599-600). A pre-
emptive construct can be general although it cannot be comprehensive. For example, if a client in therapy discriminates between people only in terms of whether they are men or women and regards all men as alike and all women as alike, the construct men-women is a highly general one for that client, though not a comprehensive one. If during therapy he comes to make more differentiations among men and among women, but still uses the construct men-women as an important aspect of his discriminations between people, the construct would then be regarded as more comprehensive than originally, although it would not necessarily be any more general.

So far generality has been discussed in relation to certain other variables within the psychology of personal constructs. It may now be compared with a few variables in other systems. Because of the similar sound of the two words it might be assumed that generality is another term for generalization as that word was used by Pavlov and as it has been used in most stimulus-response learning theories except that of Razran. Actually they are not considered to be the same. (1) The two terms are used in different types of theory and therefore in reference to different types of variables. Generalization is used to refer to events conceived to be divided into stimuli and responses. Generality refers to events that are conceived
in terms of responses. That is, generalization is part of a psychology denoted by an equation of the type \( R - f(S) \), while generality is part of a psychology denoted by an equation of the type \( R_1 - f(R_2) \). (2) The two terms are used to refer to different parts of the total sequence of events. In the case of stimulus generalization, for example, generalization refers to the first part, \( i.e. \), to the stimuli which have varying differences from the original stimulus; generality refers to the second part, \( i.e. \), to the response which is common to several "stimuli". (3) Generalization involves the direct substitution of one stimulus for another and the elicitation of nearly identical responses in the two situations. In generality of constructs the response to the new event is not a facsimile of the response to the familiar event, but is a conceptualization which includes an abstraction of similarity to at least one other event and difference from at least one other event. (4) When several new stimuli are presented in succession, generalization involves a separate response to each stimulus. Generality of constructs may involve an organization of responses in the form of a conceptualization which includes all of the events. (5) Generalization, as usually used, involves a single or relatively simple new stimulus in place of a simple old one. In generality of constructs, both old and new events usually are
quite complex and extend over a relatively long period of time. (6) Generalization occurs in terms of a gradient (often conceived to be a continuous, quantitative one) of similarity between the old and new stimuli. On the other hand, there is no gradient of generality.

(7) As used by most experimenters, generalization is a function of "objective" similarities between stimuli (i.e., similarities as defined by the experimenter) and the ability of the species, to which the subject belongs, to react to a particular variable. Generality of constructs is a function of the perceived similarities and differences (i.e., as defined by the subject; there may be considerable variation from one individual subject to another) and the construct system of the individual person. A statement about the direction of generalization can be made relatively independently of any particular subject. A statement about the direction of generality must be made in reference to a particular subject and must involve a considerable knowledge of that subject's construct system as well as the events to be construed.

So far generality has been compared with generalization as it is used by Hull, etc. In a comparison of generality with Razran's "true generalization", some of the points previously mentioned still apply and some do not. Generalization, as used by Razran, and generality
differ in the following ways. (1) Generalization is used in a stimulus-response theory and generality in a response-response theory. (2) Generalization refers to the first part of the sequence of events, generality to the second part. (3) Apparently generalization involves a separate response to each stimulus. Generality may involve an organization of responses in the form of a conceptualization. (4) Generalization usually involves relatively simple stimuli, but generality usually involves complex events. (5) Generalization occurs in terms of a gradient, although a crude, qualitative one, but there is no gradient of generality. In some respects generality and Razran's generalization are more nearly alike than generality and Hull's generalization. (1.) Hull's generalization involves the direct substitution of one stimulus for another. Razran's generalization and generality both involve cognitive responses. (2) Hull's generalization is a function of "objective" similarities between stimuli. Razran's generalization and generality are functions of perceived similarities and are not necessarily closely related to the physical dimensions of the stimuli. (3) Hull's generalization is said to occur at the time of presentation of the original stimulus. Razran's generalization is said to occur at the time of presentation of the new stimulus. Likewise, the conceptualization necessary for generality
cannot occur before the occurrence of the new event. Kluver describes generalization as a form of equivalence of stimuli or equivalence of responses. Generality of constructs appears to be closer to his equivalence of stimuli than to stimulus generalization, at least as generalization is used by all except Razran. However, generality is not regarded as the same as equivalence of stimuli. (1) Kluver points out that in order to understand equivalence one must also examine non-equivalent situations and look for something accounting for the difference between the equivalent and non-equivalent situations. The attention to difference factors is more explicit than in the case of generalization as it is usually treated. Generality of constructs also involves explicit emphasis not only on likeness but also on contrast factors. (2) Kluver stresses similarity of stimuli as perceived by some organism, whereas most users of the generalization concept have been concerned with so-called objective similarity of stimuli. Generality of constructs likewise involves perceived similarities. (3) Kluver points out that the same stimulus may elicit widely different responses and that the same response may be elicited by widely different stimuli. Although users of the term generalization have recognized that the response to a given stimulus is not always the same, they have emphasized relative similarity of
the responses. The concept of generality of constructs also involves an expectation of widely different overt motor or verbal responses to the same event, with variation from time to time and person to person. Despite the greater similarity between equivalence and generality than between generalization and generality, there are also some differences. (1) The two terms equivalence of stimuli and generality of constructs are used in reference to different aspects of the total situation. In equivalence of stimuli emphasis is still largely on the stimulus aspect, or first part, of the total sequence of events; in generality of constructs emphasis is on the second part of the total sequence. (2) Equivalence of stimuli and of responses is used in reference to situations analyzed specifically in terms of perception and motor response. Generality of constructs involves conceptualization as an important aspect. (3) Equivalence involves a relatively separate response to each stimulus when several stimuli are presented in succession. Generality of constructs often involves a good deal of organization of responses to successive events, or even organization of responses to events separated by considerable periods of time, and the final response may be unlike any of the earlier responses.

Sometimes instead of referring to the empirical phenomenon of equivalence of stimuli the term generalization is
used to refer to "going from the particular to the universal" (77:335), as Mowrer, for instance, has used it. Except for the word universal, Mowrer's definition is like that commonly used in the English language. That is, generalization, as commonly used, refers to the derivation of over-all principles from particular examples. When the word is used in this sense, generalization, along with abstraction, is conceived to be one of the processes involved in the formation of constructs. Generality, on the other hand, refers to the range of elements included in the constructs. Generality may be thought of as a product rather than a process.

The generality of a construct is analogous to Rapaport's use of the phrase realm of a concept. He states:

"Every concept has a realm: these are all the objects that are subsumed under that concept by virtue of having in common its content. Thus, for instance, the realm of the concept 'table' is all the tables which exist or can be thought of, irrespective of their material, shape, color, number of legs, or use. The content of the concept 'table', which may be referred to as 'tableness', is the elusive common characteristic of all tables." (89:389)

Presumably Rapaport would use the phrase "size of a realm" to refer to what is here called "degree of generality". Although they are analogous, generality and Rapaport's realm are not considered to be identical because of the difference in the use of the words construct and concept. The word concept, for him, refers to the likeness aspect
of all of the objects subsumed by the concept. This is shown in his statement that the content of a concept is "... the sum total of all of the characteristics which are common to all the objects subsumed under that concept" (89: 389). The word construct, as used here, gives explicit emphasis not only to the way all the objects or events are alike, but also to the way certain other objects are different. That is, a concept might be considered to be a category, while a construct is considered to be a dimension with two end-points.

Generality of Projective Test Data

An assumption commonly made when psychological tests are used is that the data obtained from the tests will enable the psychologist to make predictions about the subject when he is in other, non-test situations. This involves the generality of the subject's responses to the tests. That is, the question may be asked whether the responses to the tests are broadly characteristic of the subject, or whether they are specific to the particular tests used. The assumption is nearly always made that they are not specific to the particular tests. This assumption is especially shaky in regard to many so-called projective tests, which have become increasingly more popular in the past decade, for the content of the responses to many
of them is far removed from non-test behavior.

The terms projective technique and projection have been used variously with some resultant confusion. The word projection was introduced by Freud in 1894 in a paper on "The Anxiety Neurosis" and was amplified by him in later papers (1:7). A definition by Sears describes the most common use of the term by both psychoanalysts and non-psychoanalysts.

"A wish, attitude, or habit-hierarchy which is not compatible with other attitude or habits of an individual may be attributed by that individual to other persons rather than to himself, providing he lacks insight into the fact that he himself possesses the trait in question. This process of attribution is unconscious, i.e., the subject does not give any verbal evidence that he knows his perception is false." (104:561)

A brief history of some of the highlights in the development of this concept of projection is given by Abt and Bellak in their new book (1:7-11). Sears (105) distinguishes this concept from motivationally determined perception, which has also been called projection. In line with the latter usage, there have been a good many demonstrations of the influence of such factors as fear (79) and hunger (99, 100) on perception.

Although the most common use of the word projection refers to a paranoid attribution of traits to others, this is not the most frequent usage of the word in the phrase "projective technique". Frank first used the phrase pro-
jective method in 1939. He explained his meaning by stating that

"A projective method for the study of personality involves the presentation of a stimulus situation designed or chosen because it will mean to the subject not what the experimenter has arbitrarily decided it should mean (as in most psychological experiments using standardized stimuli in order to be 'objective') but rather whatever it must mean to the personality who gives it, or imposes upon it, his private, idiosyncratic meaning and organization." (27)

This use of the term projective is more akin to Sears' motivated perception than it is to the paranoid type of projection. Closely related to Frank's concept is the idea of individual expressiveness, which was stressed by Rapaport in his "projective hypothesis" when he said that "All behavior manifestations of the human being, including the least and the most significant, are revealing and expressive of his personality, by which we mean that individual principle of which he is the carrier." (88a)

In an attempt to clarify the concepts further and to relate projection and projective techniques, Cattell (17) stated that the responses to a projective technique show three major types of projection. The first of these, which he terms N.I.P.E. projection, is "projection through naive inference from limited personal experience". Here the subject "deduces ... that the behavior of the character he is watching springs from a certain motive, A, which he himself would have if he were presented with the same situ-
The second type, "unconscious, immediate, or true projection", is the type that Freud spoke about. It involves attribution of an unacceptable drive or desire from within the subject to another person. The third type, P.R.E.S. projection, is "projection of press required by emotional state". In this type there is distortion of perception so as to make the world fit in better with the subject's prevailing emotions. This is the type most like motivated perception. If people think of projection only in terms of the second type (and many do this), then the name projective technique is a misnomer. Bellak has proposed that apperception be substituted for projection in the name of those tests because they involve "... an organism's dynamically meaningful interpretation" (1:11-12). It is in the sense of a meaningful interpretation that the term projective technique will be used within the psychology of personal constructs.

Generality as a factor in projection and projective techniques as a whole has been of concern to Cattell and to MacFarlane and Tuddenham. In addition it has been important in validity studies of specific techniques. Wyatt (136) and Wittenborn (131) for instance, have considered it in regard to the Thematic Apperception Test.

Speaking within a psycho-analytically influenced framework, Cattell has summarized his views on the generality of
the three types of projection described earlier.

"Is projection a general process, likely to operate in each and every real life or test situation we like to set up, or will it be a specific to one restricted life situation? If a patient is handling a powerful, frustrated and repressed aggression by projecting it on his father, is that not in itself sufficiently cathartic and adjustive to free him of any necessity to project aggression on every individual he rates and every projection test fictional person he describes? An answer is possible with respect to two of the three kinds of projection. The N.I.P.E. projection of conscious feelings will be specific to the situation; e.g., if he hates his father, he will assume that the other children hate their fathers, but not that every person is hateful. The P.R.E.S. projection in response to free-floating emotion will presumably tend to be general; e.g., a negative self-feeling or inferiority will tend, as Sears has shown, to issue in a belief that practically anyone is very critical.

"No answer can confidently be put forward with regard to true projection, but an intermediate degree of specificity is indicated. True projection is generally observed clinically not as projection of single traits but of a dynamic situation, plot, or thema. The individual projects not only the drives of the id but also the opposing drives of the super-ego and the anxiety of the ego between them. We should expect therefore that the subject would project with different degrees of power into different situations (for an unknown person is a standard peg on which almost any motive may be hung, but a structured situation, even when vague, will not fit every thema), corresponding to their convenience to his thema." (17)

Most users of projective techniques have probably not stopped to think of projection or its generality in terms of the three types described by Cattell.

In accordance with the meaning of "projective techniques" as used by Frank and others, MacFarlane & Tuddenham have pointed out some of the assumptions that are us-
ually not recognized, at least explicitly. The basic assumption -- which usually is recognized -- is that "... every subject's responses are not the consequence of sheer accident but are determined by psychological attributes of that subject" (6:33). This assumption involves three corollaries, and it is these corollaries which are usually not stated. These corollaries may be listed as follows:

1. "... belief that a protocol is a sufficiently extensive sampling of the subject's personality to warrant formulating judgments about it."
2. "... belief that the psychological determinants of each and every response are basic and general."
3. "... belief that projective tests tap the durable essence of personality equally in different individuals."

They state that the first of these corollaries remains to be demonstrated, but that the second and third have been contradicted by evidence. In regard to the first corollary they point out that "There is no satisfactory evidence on the general question of how many responses must be secured from a subject to justify confidence in the stability of the various scores arrived at in analyzing a protocol."

Macfarlane herself demonstrated the lack of validity of the third assumption in an investigation of boys and girls as a part of the California Child Guidance Study. In regard to corollary number 2 it appears to the writer that until recently many users of projective methods dis
assume that all responses are equally revealing, but that in the past few years there has been increasing recognition that this is not so. The corollary has been recognized and rejected by those who have investigated the influence of the structure of the test materials upon perceptions and responses. In regard to the Thematic Apperception Test, for example, this has been of concern to Eron (24, 25), Garfield and Eron (28), Rosenzweig (94), Rosenzweig and Fleming (95), Markmann (68), and Lindzey (65).

One of the most popular projective methods involves the use of pictures along with instructions for the subject to make up stories about the pictures. A history of the development of this technique up to 1947 is given by Tomkins in his book The Thematic Apperception Test (119:1-20). Probably the first use of this method was in a 1907 investigation by Brittain, comparing the imagination of adolescent boys and girls (14). Twenty-five years later, in 1932, Schwarz(103) used Brittain's technique as an aid in psychiatric interviews with delinquent boys. Using pictures representative of situations familiar in the life histories of delinquents, he asked the boys to describe what the pictured boys were thinking and then to tell what they themselves would think and do if they were the boys in the
pictures. In 1935 Morgan and Murray first published ma-
terial about the use of the picture-story method in the
investigation of fantasy in college students (76). Unlike
Schwartz, they did not ask the subjects to describe the
thoughts and actions of both the pictured figures and
themselves, but aimed to distract the attention of the sub-
jects from themselves. The use of this method, they said,

"... is based on the well recognized fact that when
someone attempts to interpret a complex social sit-
uation he is apt to tell as much about himself as
he is about the phenomena on which attention is fo-
cused. At such times the person is often off his
guard, since he believes that he is merely explai-
ing objective occurrences. To one with 'double
hearing', however, he is exposing certain inner
forces and arrangements - wishes, fears, and traces
of past experience." (76)

Murray has published several sets of pictures, and in
1943 the set currently in popular use appeared (80). Since
then a growing number of other sets of pictures have been
published. Some of them have been modifications of the
Murray pictures for the study of special groups; thus
Thompson (120) has substituted Negro figures for the white
ones in the Murray situations. More often the sets have
used entirely new pictures and they have often been inten-
ded for use with specified groups of subjects or for spe-
cified but limited purposes. Symonds (117, 118) has de-
vised pictures for adolescents; Henry (42), studied various
Indian groups with pictures drawn to represent situations
familiar in the lives of Indians; Bellak and Bellak (10)
have published a set of pictures for children from 3 to 11 years old; extensive research is being done on the Michigan Picture Test (27) for children 8 to 14 years old; Blum (12, 13) developed the Blacky Pictures to study psychosexual development according to psychoanalytic theory. This list represents the majority of sets of published pictures, although it by no means includes all the sets that have been used in research. Wide variations will be found in the care with which the pictures have been developed and in the amount of work that has been done on standardization of responses to them. The most thorough development is probably being done in connection with the Michigan Picture Test. No doubt the sets will also vary considerably in their clinical usefulness. For instance, doctoral research by Mitchell (75) has suggested that at least for northern Negroes the Thompson modification is no more useful than the Murray pictures.

Many methods of analysis of picture-story protocols have been devised. Nearly all of them were worked out with reference to Murray's Thematic Apperception Test pictures, but they could be used with other sets of pictures as well. Shneidman's new book, *Thematic Test Analysis* (108), presents sixteen methods of analysis by fifteen authors, and there are at least a half dozen additional methods not reported there.

An organized survey of methods of analysis could be
made from any of several starting points. Several possible divisions of methods will be described and illustrative references will be cited. Methods might be divided into two groups on the basis of use of some sort of scoring method (e.g., 1, 36; 119) or use of interpretive hypotheses without a formal scoring system (e.g., 97, 97a, 118; 6: 181-229). They could be divided into a larger number of groups on the basis of the type of theory or kind of assumptions made by the authors. For instance, Murray's need-press system has been used in its original or modified form by Murray (80), Aron (7), and Cox & Sargent (2). Psychoanalytic theory, with varying emphasis on the use of symbolism, has been followed by Masserman & Balken (72), Holt (6: 181-229), and Aron (7). A normative approach is emphasized by Eron (24), and is also used as part of the analysis of protocols by Rötter (97, 97a), Wyatt (135), and Harrison (35).

Another type of analysis of methods is in terms of aspects of the protocols which are emphasized. Here any division is likely to be artificial, because all of the authors utilize several aspects. However, there is variation in the combination of aspects emphasized and in the degree of relative emphasis of different aspects. Three broad areas of emphasis may be distinguished: story content, formal characteristics of the protocols, and figure analysis. In the
category of story content, need-press themes have been emphasized by Murray (80), Klebanoff (108:126-131), etc. Frequency and unusualness of themes are used by Rotter (97, 97a), Harrison (35), and Symonds (118), and type of themes is used by Eron (24). Types of outcome or consistency of outcome with the rest of the story has been utilized in many methods (e.g., 24, 97, 97a, 95:139-159, 1). Types of conflicts or worries, methods of solving problems, the handling of blame and punishment, and sequence of events are among other areas of content analysis that are investigated by various authors. Formal aspects of the stories are also analyzed along many lines. Among them are the degree to which the subject follows the test directions (135, 24); logicalness and coherence of stories (97, 97a); choice of language, and style (35, 135); level of interpretation (119, 95); misperceptions, omissions, or additions of objects (1, 24); and emotional tone or mood of the story (89, 24, 28). The amount and type of emphasis on the analysis of the pictured figures has varied, perhaps more widely than the emphasis on either story content or formal characteristics of the stories. Many authors look for the hero, or figure with which the subject is supposed to identify. They may make the analysis in terms of feelings and thoughts along (72) or in terms of activities (135). They may attend chiefly to the hero (119) or they may consider the hero of
little more importance than the other figures (85). They may examine other figures for the hero's attitudes toward them (Holt in 6), they may consider the other figures to represent people important in the subject's life (89), or they may look for projection of the subject's unacceptable characteristics onto the other figures (85). Or, instead of characteristics of the individual figures, personal relationships and interactions may be stressed (Arnold in 108).

From the many sets of published pictures and the numerous variations in methods of analysis of protocols one gets a picture of a bewildering range of possibilities for investigating the verbal behavior of either clinic patients or college freshmen. But despite the varying methods of analysis and sets of pictures, there are at least two features common to all or nearly all users of this type of test.

A feature common to all the authors of sets of pictures and of methods of analysis is an attempt to abstract from the protocols characteristics of the subjects that are sufficiently widespread as to be of functional importance in other situations. Since their theoretical approaches and assumptions vary, they naturally differ in regard to what aspects of the subjects they believe are important in both the test and daily life. Murray (80) believes that the test
shows "ideas, plans, fantasies and dreams about behavior" rather than primarily overt motor and verbal behavior. He makes two assumptions: (1) "... the attributes of the heroes (needs, emotional states and sentiments) represent tendencies in the subject's personality" and (2) "...the press variables represent forces in the subject's apperceived environment, past, present or future". A similar view is that of Bellak (1), who holds that the subject's feelings, sentiments, and drives are common to both the test protocol and everyday situations. Lindzey (65) has stated that the primary assumption involved in the use of the tests is that the subjects will reveal their "strivings, dispositions, and conflicts". Although Piotrowski realizes that in one case the subjects are reacting to pictured figures and in the other case they are interacting with real people, he still believes that often the protocols reflect daily overt conduct (65).

A feature common to many of the authors, especially in the more recent publications, is a recognition that the test responses are a function not only of some aspect of the subjects themselves but also of the structure of the pictures and the society in which the subjects have been reared. Recognition of the influence of picture structure upon the responses of subjects has made Rosenzweig state
"Since the subject is permitted to define or organize the partially structured stimulus materials according to his inclinations or interests he is thus both 'externally' determined and 'internally' free. He may respond as he wishes but it now becomes necessary for the examiner to decide how far the structure of the stimulus has affected the response and how far some personal idiom has prevailed... it is possible to evaluate the external or internal determination of the subject's behavior only by consulting an empirical catalogue of responses elicited from other subjects who are comparable to him in age, sex, etc." (94)

Even before Rosenzweig's statement appeared, the same recognition had already led to the use of informal norms for the usualness of responses by such men as Rotter (97).

In the last three or four years normative material has been published on limited numbers and kinds of subjects by Eron (24,25), Rosenzweig and Fleming (95), Bellak, Ekstein, and Braverman (11), and Cox and Sargent (21).

The importance of the structure of the materials in determining responses to the pictures means that a limit is placed on the assumption of broad generality of characteristics shown by the subjects on the test. It is possible that some, or even a large proportion of responses, may be socially conventional reactions to specific situations and that they are common to many people rather than characteristic of any one individual. Light has been shed on the problem by research in which comparisons have been made between test variables and
some sort of behavior in other situations, i.e., by validity studies. It has been seen that many different sets of pictures have been devised. The importance of picture structure in determining responses means that new investigations must be done with each set of cards. To date, most of the investigations have used the Murray cards.

The problems and findings in validity studies up to 1947 have been reviewed thoroughly by Tomkins (119:9-19). Holt (6:221-223) has done a briefer survey, but cites material published up to 1951. Because many of the users of this type test assume that the responses are more relevant to the subject's thoughts, wishes, and fantasies than they are to his overt behavior, there have been comparisons of Thematic Apperception Test responses with dreams (101) and with material revealed during psychoanalytic sessions (76). The results, which were not analyzed quantitatively, but on a case study basis, were considered to show satisfactory agreement between the themes revealed by the two methods. Biographic and autobiographic materials have been used as a criterion by several psychologists. Harrison (33,34) found good validity indices when inferences drawn from the TAT were checked with biographical and personality items from case histories of psychiatric hospital patients, and there was
not much difference between the two indices. Combs (20, 20a), using a questionable method, found unusually high agreement between goals and desires as expressed in the TAT and in autobiographies. Markmann (68) found that validity was specific to particular themes elicited by certain pictures. Similarity of findings from the Thematic Apperception Test and other tests, particularly the Rorschach test, has been reported by Harrison (35) and by Henry (42). Hartman (36) reports statistically significant correlations between TAT variables and a combination of interview and test data, but his correlations are so low that they are not likely to be important when applied to individuals. Groups differentiated in various ways have been studied, and differences in TAT results have been reported. Cox and Sargent (21) found that, in comparison with emotionally stable boys, emotionally disturbed boys gave shorter stories, more often expressed no feeling, gave fewer stories of threats of disaster, death, and domination, more often gave stories without action or without endings, and more often failed to identify with a hero. Balken and Masserman (7a) have found differences between groups varying in psychiatric diagnosis, and Renaud (91) found differences between groups having brain disease, brain injury, and no brain pathology. Butler (15) failed to find much between-group difference in the handling of
parent figures by dependent children in foster homes and in an institution, but she did find within-group differences; she did not attempt to validate the within-group differences. Richardson (92) found no significant differences between stutterers and non-stutterers on the TAT, although he did find differences on other tests. Predictions of changes in protocols as a function of experimentally controlled situations have been made by Bellak (9) who found an increase in aggression after the subjects had been criticized, and by Rodnick and Klebanoff (93), who found a decrease in themes of superiority and an increase in aggression after induced frustration. In an investigation done within the framework of Rotter's Social Learning Theory of Personality, Crandall (22) demonstrated a decreased freedom of movement following frustration, and the area of decrease showed some relationship to the area of frustration. Bellak (18-10) also found changes corresponding to posthypnotic suggestions to feel aggression, sadness, or elation.

All of these studies have used interpretations of test protocols instead of the protocols themselves and nearly all have also involved a high degree of interpretation or judgment in the criteria. Harrison's comparison of TAT interpretations with biographic details involves the smallest degree of interpretation of the criterion. Because
of this, the results are valid only for the particular method of interpretation used and perhaps even for the particular interpreter. This is especially important in view of the large number of methods of interpretation that have been described. Whenever interpretations involve much abstraction from the data, there is increased opportunity for error, and the chances of biased judgments are multiplied when both the test data and the criterion are abstracted. This is true particularly if the abstractions are on the level of wishes, thoughts, or fantasies, rather than of overt behavior. For determining the applicability of test results to other situations, it is desirable to reduce the degree of abstraction to a minimum and compare, as far as possible, overt behavior rather than thoughts or fantasies, in relation to both the test and criterion situations. The plan of the present research is to compare two forms of language behavior in a way that avoids almost all abstraction by the experimenter.

So far in this section the discussion has been about projective techniques entirely. The question might be asked how they are related to the psychology of personal constructs.

The psychologist working in a practical situation is often called upon to make predictions about the future behavior of people in a variety of situations. In a mental
hygiene setting, for example, the decision must often be made as to whether to offer a client counseling; this involves prediction of his behavior both if treatment is given and in the event it is not given. Another problem may be one of obtaining some notion of the areas of difficulty before therapy or early in the therapy. Or the problem may require predicting behavior if any one of several alternative types of treatment is used. According to the psychology of personal constructs the accuracy of such predictions is likely to be increased if the psychologist has some knowledge of the client's expectations of other people and of the pathways along which he is free to move. The expectations and pathways are believed to be shown by the constructs that the client uses to describe people and events around him. It is held that as the client has participated in or observed events he has formed constructs about those events, and that the constructs are directed toward optimal anticipation of future events. In construing people he has been predicting what those people will do. His predictions or expectations of the actions of other people, in turn, help to determine his own actions. Thus a knowledge of his major constructs can aid the psychologist in predicting the client's behavior in the future. The two-ended quality of constructs also enables the
psychologist to anticipate the most likely direction of first changes in behavior in case some events should close his present paths of action. The most available alternative is held to be the direction represented by the contrast end of a construct that is governing present behavior. But if the psychologist is to make use of the client's constructs, he must have some way of obtaining a sample of them. It is held that tests are useful in getting such a sample.

In his manuscript Dr. Kelly has discussed the functions of tests in some detail. He has stated that "... the primary purpose of psychological measurement in a clinical setting is to survey the pathways along which the subject is free to move" (53:124). As has been stated, these paths are believed to be shown by the subject's constructs. Another important function of tests is to define the client's problems in a form that the psychologist can use. A knowledge of his constructs can help in this task, also. For example, if he construes the majority of people as either "kind, understanding, and sympathetic" or "not interested in me", we get a picture of a person who is childishly self-centered, whose major interest in other people is of a receptive type, who is lacking in much understanding of other people, and who forms sweeping opinions. A third function of tests is to furnish testable hypotheses about the client; this is in contrast to the use of tests, in
some quarters, as microscopes for the examination of the inner being of a patient. In regard to the hypothetical client mentioned above, the construct "sympathetic - not interested in me", along with other constructs, might lead to a hypothesis that the client is likely to be demanding of other people but unable to give much to others except, perhaps, as bribes for sympathy. A fourth function of a test is to show the psychologist resources of the client which may be used in treatment or in any type of readjustment. For instance, if the client's test protocol shows minor constructs of warm-cold, generous-stingy, and good looking--not good looking, it might be possible, in therapy, to bring about a restructuring of the constructs so that the client comes to include himself among the warm people and the generous people; then perhaps he can interact with other people more on a give-and-take basis and less on a receptive, demanding basis.

"Projective" techniques, as distinguished from other types of test, are held to reveal the client's own constructs rather than to investigate his ability to use the constructs of the examiner (53:187). The importance of knowing something of the client's own construct system is self-evident, if one holds that behavior is in terms of personal constructs. To be most useful the test methods should elicit constructs of the kind that the client is likely to use in everyday situations. That is, the constructs
elicited should have enough generality to appear in non-test situations. Obtaining a survey, through tests, of constructs related almost exclusively to inkblots or to pictures of strange people (e.g., psychiatric hospital patients) is a waste of time in the handling of most clients. On the other hand, tests which elicit constructs about the kinds of people that the subject may meet are likely to be particularly useful because interaction with other people is an important part of the lives of most clients.

In order to bring out a client's constructs about real people Kelly has worked out the Repertory Test (53: 127-185), which forces the subject to construe people he knows. After the subject has made up a list of names of relatives, friends, and associates or acquaintances known in various types of situation, according to certain specifications given by the examiner, he is presented names of three people at a time and is requested to tell in what way two of the people are similar and yet different from the third person. Because the test makes use of people known to the subject, it is likely to bring out constructs whose range of generality includes people met in non-test situations. However, there are likely to be clients with whom it is desirable to use a more indirect technique. Picture tests look quite promising for eliciting constructs. In
telling stories and describing pictured people the subject may reveal the dimensions he uses to construe events and therefore the pathways along which he himself is free to move. He may show his views of the kind of world he believes he must adjust to, fight, or try to change, and he may also show the methods that are open to him. On the one hand it is recognized that the generality of the constructs elicited may be limited by the fact that pictured people rather than real people are construed. On the other hand, a naive subject may feel more freedom to express socially disapproved constructs while talking about pictured people than would be the case if he were describing real people known to him. The present research is concerned with the generality of constructs about people elicited by a version of the Repertory test and a modification of the Thematic Apperception Test. This topic will be discussed more fully in the following chapters.

Emphasis on the generality of constructs elicited by a picture test is closely related to validity studies of picture tests. Both emphasize the relatedness of two samples of behavior or two forms of data. If constructs used on a picture test are demonstrated to have some generality to other situations, this may be regarded as a demonstration of validity of a particular type. From the studies reported on pages 65 to 67 it can be seen that although some relationships have been shown between Thematic
Apperception Test Stories and other data, none of the studies has been concerned with the language or constructs used in the test protocols and in other situations. It is held, within the psychology of personal constructs, that such an approach is a useful one.
CHAPTER III

STATEMENT OF THE PROBLEM

Origin of the Problem

The writer's interest in the generality of projective test protocol arose several years ago in a clinical setting. At that time some users of the Thematic Apperception Test were making sweeping claims about the powers of those methods, with little experimental foundation for their statements. They often implied extremely wide generality of whatever the tests were supposed to measure. The writer felt dissatisfied with some of the claims that were being made and looked about for more reasonable and better validated methods of interpretation.

Use of the techniques within the framework of Freudian or Rankian psychoanalysis (the theoretical orientation of the non-psychologists on the clinic staff) involved adoption of all the defects in those theories. Heterogeneous brands of learning psychology based on rats or nonsense syllables (the views current among most of the psychologists on the staff) offered little help. Most of the published methods of analysis of the Thematic Apperception Test required the assumption of identification by the subject with a hero figure (cf. Murray(80,81) and Tomkins (119)). Not only was
this assumption questionable, but it also led to problems in the discovery of the hero and the reconciliation of contradictory heroes. Murray's need-press system and Tomkins' analysis of vectors, levels, conditions, and qualifiers were too unwieldy to be practical for regular use except for research. At that time most of the interpretation of the TAT was based on analysis of plots or themas, while relatively little use was being made of the figures themselves or the structural characteristics of the stories. Despite the use of the plots of the stories by some people in a literal sense to indicate how the patients were likely to behave in non-test situations, it was obvious that they should not be used in this way. The range of situations was limited and the patients were instructed to use their imaginations and to make up stories. Some of us at the clinic were impressed by the frequent recurrence of particular plots, and we began to wonder whether the plots did not tell more about culturally common reactions to the particular pictured situations than they told about the individual patients. At that time the normative studies of Eron (24,25,26) and Rosenzweig (94,95) had not yet been published, but they subsequently confirmed the impression that certain plots are characteristic of particular pictures in the Murray set.

In our search for a more valid method of interpretation, some of us began to make increasing use of aspects of the
protocols other than the plots. For instance, we began to abstract from the stories descriptions of the people in them in hopes of being able to make more accurate predictions of how the patients saw people in their environment and consequently how they might act. But we still wondered what we were getting. Were the abstractions about the people as much a function of the particular pictures as the plots seemed to be? Did the patients actually see people around them in the ways that they used to describe the story character? Could their views of people be abstracted efficiently in terms of any particular groups, such as age-sex groups? All these questions remained unanswered at that time except for the impressions obtained from the protocols and from case histories and interviews with patients.

Several years after these questions about the generality of projective test protocol arose, the writer became familiar with the psychology of personal constructs as it was being developed by Dr. G.A. Kelly. The emphasis on behavior in terms of personal constructs seemed to "make sense". It offered a way of dealing with individuals instead of with group trends, it showed a way of working with large segments of behavior rather than small units, it gave a way of handling social relations and not merely conditioned responses or memory tasks, and it offered a method of predicting behavior rather than merely explaining it in postdictive fashion.
Constructs about people and interpersonal relations interested the writer particularly, because of the large social element in our lives. She began to wonder about possible patterns for construing people and about the range of use of constructs. Are a person's constructs sufficiently personal that they can be distinguished from the constructs of other people? That is, does a person's construct system have enough inner coherence or consistency as well as enough difference from the systems of other people to be distinguished from the systems of others? As a rule does a person tend to use the same set of constructs in all or most types of interpersonal situations, or does he have somewhat different sets for different types of situations. Since in our culture many events are related to the age and sex of people, are a person's constructs about others likely to be differentiated into groups, even if overlapping groups, on the basis of age or sex or both? Are constructs elicited by psychological tests likely to be used in other situations, or are they limited largely to the tests? Dr. Kelly has been using the Repertory Test and self-characterization sketches to obtain subjects' constructs about themselves and other people. Could the Thematic Apperception Test pictures also be used effectively to sample a subject's constructs about real people? These are some of the questions in regard to personal constructs about people that come to mind. It can
be seen that the old questions about the range of applicability of material obtained from projective techniques had been absorbed into the newer interest in personal constructs.

Implications of generality in the Psychology of Personal Constructs.

Any system of psychology that tries to explain and predict behavior must deal in terms of repetition of perceptibly similar occurrences. This does not mean that it is necessary for any event ever to be exactly like any other event. It does mean that events can be handled only if people at least view them as somewhat like other events. It is hard to conceive of any system for effectively organizing behavior if every event is seen as completely different from every other event, or even if every event is seen as completely different from all but, say, two other events.

Now a construct has been defined as a way in which two things are alike and yet different from a third thing. A construct, therefore, is a perceived regularity between at least three things. If a construct subsumes a large number of elements, the person using that construct sees some regularity among a large number of events. And if constructs involve regularities among a relatively large number of events, it is possible that the constructs may be used for organizing behavior. But if each construct involves regularity among the
minimum number of events (three), constructs are likely to be of little use in organizing behavior, and a system of psychology based on them will show little relationship to the events of life.

This situation can be expressed, also, in terms of the generality of constructs. It will be remembered that the generality of a construct is expressed by the elements which that construct subsumes, and that the degree of generality refers to the number of elements subsumed by a construct. If a construct subsumes a certain number of elements, i.e., has a certain degree of generality, the person using that construct sees some regularity between those elements. Now if an important number of a person's constructs have an appreciable degree of generality, it is possible that the constructs may be used to organize behavior. But if all of an individual's constructs have a low degree of generality, then use of them for organizing behavior will be ineffective. The utility of a system of psychology based on constructs depends on an important number of constructs having a significant degree of generality.

But now take the case of a system of psychology built not merely on constructs, but on personal constructs. Personal constructs may be distinguished from public constructs on the basis of use by a single individual as contrasted with use by a relatively large number of individuals. It should be remembered that use of the same words is not the criterion
of use of the same construct. Several people may use the same verbal symbol, but refer to different constructs because the elements grouped together to form the construct are different. In the case of a public construct, many people agree on the elements subsumed by the construct. "Table" may be considered a public construct in our culture. A private construct involves considerable difference from person to person in regard to the elements subsumed by the construct; difference, as used here, means that on the level of pointing to people, objects, and events subsumed by a construct there is disagreement among people. An example of a private construct might be the one symbolized by the word "friendly". Even though the word "friendly" is used by many people, the construct may be considered a private one because people vary in regard to which people and which behaviors they consider friendly. The utility of a system of psychology built on personal constructs, like one built on public constructs, is dependent on a considerable number of the constructs having an important degree of generality. In addition, a system built on personal constructs requires that the constructs of various people be differentiatable. If there are not detectable differences between the constructs of various people, then the psychologist might as well act as if all constructs are public.

Generality and distinguishability of personal constructs might be assumed by P.P.C. However, it is possible to de-
monstrate both and so remove them from the area of assumption. Such a demonstration is part of the present research. The present research combines theoretical and clinical purposes through a test of some aspects of the generality of constructs obtained from certain projective techniques.

Exploratory Research

During the summer of 1951 free research was carried on with the aim of setting up a fixed design usable for investigating the generality of personal constructs. At this early stage the plan was to investigate generality as it might be shown in several samples of verbal behavior. It was felt that the larger the number of samples from each person, the greater would be the chances of demonstrating whatever generality might be present. Feasibility, however, appeared to dictate that the number of samples from each person should be limited to three. Three tests were chosen which would all elicit constructs about people: A selection of pictures from the Murray and Symonds sets, Kelly's Repertory Test (53:127-185), and Kelly's Self-characterization Sketch (53:306-334). It was expected that three protocols by a single subject would show a good deal of overlap in the constructs used. In other words, it was thought that the generality of the majority of constructs used by any one subject would be wide enough to include elements that would appear in at least two of the three verbal samples. A preliminary hypothesis
was set up that, on the average, three protocols by a single subject would show more similar than dissimilar constructs. The purpose of this early research was threefold: (1) to work out methods for the use of the three tests; (2) to set up a scheme for scoring similarity of constructs; and (3) to obtain preliminary results of a test of the hypothesis in order to be able to refine the hypothesis.

For the picture-story test eight pictures were selected from the Murray set and 4 were selected from the Symonds set. Men were given Murray pictures 10, 7BM, 4, 6BM, 14, and 6GF plus Symonds pictures A2 and A7. Women were given Murray pictures 10, 7GF, 4, 9GF, 14, and 6GF plus Symonds pictures A5 and B8. Criteria for the selection of pictures included: (1) Avoidance of pictures eliciting highly stereotyped plots or characterizations, as determined by subjective impression from the writer's clinical experience; (2) For each sex of subject, inclusion of at least one picture with a young figure of the same sex (as the subject) alone; one with a young same-sex figure with opposite-sex peer figure; one with a young, same-sex figure with an older figure of the same sex; and one with a young, same-sex figure with an older figure of the opposite sex (No suitable pictures of 2 young figures of the same sex were available); (3) For all subjects, inclusion of one picture with a figure of indefinite age and sex alone, and one picture with two indefinite figures together. Eron & Ritter (26) had found that although oral and written ad-
ministration elicited similar themes, they showed marked
differences in the formal aspects of the stories. There­
fore the test was administered orally to part of the sub­
jects and in written form to part of the subjects in order
to determine relative effectiveness of the two methods in
eliciting descriptions of the people.

Three forms of Kelly's Repertory Test were used. In
this test subjects are given a list of "role titles" and
are instructed to write down names of people known to them
who fit the titles. The titles in the list may be varied,
but such titles are used as "a teacher you liked", "your
wife or present girl friend", "your father", "a person with
whom you have worked who was hard to understand", "the per­
son whom you would most like to help", and "the most inter­
esting person you know". During this part of the investi­
gation two different, but partially overlapping lists of
titles were used; the first list contained twenty-four titles
and the second one eighteen titles. Of the subjects who re­
ceived the first list, part were administered the test in the
Full Context Form and part were given the Minimum Context
Form; all the subjects who received the second list were given
the Minimum Context Form. The method of administration was
varied because up to that time there had not been enough re­
search with the test for the experimenter to know which form
would give a greater amount of usable data most efficiently.
The Minimum Context Form involves the presentation to the sub-
ject of sets of three names at a time along with instructions to sort the people so that two are alike in some way and yet different from the third person in that same way. Because of the activity required, each of the sets of three names is called a "sort". The first list of titles involved 32 sorts and the second list involved eighteen sorts; work with the first list of titles had shown that eighteen sorts was as many as most subjects could do in the experimental time available. For the first list the combinations of names were arbitrary, while for the second list an attempt was made to follow certain principles which included comparison of liked and disliked people, comparison of spouse and parents, comparison of siblings of the subject, comparison of parents with "ideal" older people, use of sets with titles all of one sex and sets with titles of mixed sex, and comparison of three liked or three disliked people. In the Full Context Form the subjects are given all the names at one time and are told to divide the people into as many groups as they wish in such a way that the people who are alike in some important way are put together. After the subject has completed one sorting, he can be given the full set of names as many times as the examiner wishes and experimental time allows. Use of both forms of administration with the first list of titles showed that the Minimum Context Form gave more information in the time available than did the Full Context Form, and therefore only the former was used with the second list.
Only one form of the Self-Characterization Sketch was used. In this test, if a subject were named Richard Dixon, for example, he would be instructed to "write a character sketch of Richard Dixon, just as if he were the principal character in a play. Write it as it might be written by a friend who knew you very intimately and very sympathetically, perhaps better than anyone ever really could know you. Be sure to write in the third person ..." It can be seen that the subjects are given a great deal of freedom in regard to what they may include in their sketches, and as a result the content of their sketches varies widely. They are directed to write in the third person with the hope that this form will reduce self-consciousness and therefore censorship of material.

Originally the writer had planned to administer all three tests to each subject in order to be able to compare similarity of constructs in the three records of each subject. However, such a program would require between two and three hours of each subject's time, whereas during that quarter of the academic year subjects were available for only fifty minutes each. Therefore it was necessary to limit the examination of each subject to one test, except for a single subject who was given both the self-characterization and the picture-story test. A total of thirty-eight subjects were used. This procedure meant that no comparisons could be made
between protocols by the same subjects, and so it was necessary to look elsewhere for two or more protocols by the same subjects.

It was found that during the Spring quarter of 1951 about twenty students in a survey course in clinical psychology for non-majors had completed a group form of the Repertory Test and had written self-characterizations as part of their course work. These records were examined and they were used in a little preliminary work in setting up categories for scoring similarity of constructs in the two protocols. Inspection of the protocols indicated, however, that most subjects did not show much overlap of constructs in the two records. Therefore, the attempt to set up scoring categories was discontinued.

As a result of testing the thirty-eight subjects and examining the protocols of the other group of subjects certain conclusions were drawn with reference to this and future experimentation. It was apparent that the original plan to compare three different records from the same subjects was too ambitious in terms of both the time required from the subjects and the amount of necessary analysis of the data. Therefore it was decided to compare constructs from only two records, despite a realization that two is a small sample for testing generality of constructs. As the tests were used in this free research, determination of generality was difficult or impossible because of differences in the areas of constructs
elicited by the various tests. Inspection indicated that the majority of constructs elicited by the Repertory Test might be labeled broadly "personality" or "trait" or "descriptive" constructs. In the picture-story test the majority of the material was in "action" terms, despite emphasis in the test instructions to "tell what kind of person each figure is" in addition to giving a plot. Because of the lack of structure in the directions of the Self-Characterization Sketch the type of material elicited varied widely. With such wide differences in areas of constructs elicited by the three tests it was clear that no direct comparison could be made. The alternatives were to change the method of administration of one or more of the tests or to use abstraction by the experimenter to turn the action descriptions into personality constructs. The amount of abstraction or interpretation required would introduce an opportunity for large errors, and so it was felt that a change in method of administration of the tests was preferable. A change was made which would free the subjects to do their own abstracting. Since only two tests could be used in the fixed design research, it was decided to omit the self-characterization, because that was the least structured of the three procedures. The greater amount of structure in the other two tests meant that there would be more chance of controlling the area of the constructs elicited. Another reason for retaining the Repertory Test and the picture-story test was that both
tests could be used to elicit constructs about a variety of people, whereas the self-characterization elicited material chiefly about the subject.

In regard to the comparison of responses a tentative decision was made to deal with content of the constructs rather than with any classifications according to types of constructs. This decision was based on the experimenter's interest in the applied aspect of clinical psychology and the observation that in clinics many people, both psychologists and non-psychologists, are interested in the content of the ideas expressed on tests and their similarity to the content of the patient's wishes, fears, etc., about themselves and their relatives and friends. At first after the decision was made to deal with the content of constructs, the idea of using some sort of checklist was entertained. However, two major difficulties led to the abandonment of that plan. In the first place, a checklist method would require either (1) an assumption that every use of a different word by a subject meant use of a different construct, and every use of the same word by a subject represented the same construct, or (2) the preparation of an individual "dictionary" for each subject. During the latter part of the free research period the experimenter had done some of the preliminary work of making up such dictionaries for a few subjects. For each of these subjects a list was made of the adjectives and phrases he had used in describing people on the
picture-story test; the subject was then asked to give a synonym and an antonym for each adjective or phrase. It was planned to record any relationships between the various words and phrases and to use the relationships along with the list in checking the Repertory Test protocol for presence of the same responses. In the free research period, however, such a check could not be carried out because it was impossible, to obtain both a picture-story test protocol and a Repertory Test record from the same subjects. However, the work that was done indicated that use of individual dictionaries would be cumbersome and would still involve a large chance for inaccuracy. Furthermore, there was a second difficulty in the use of a checklist. This involved the necessity of a method for equating the importance of constructs in the two protocols. An individual weighting scheme would have to be set up for each subject, and this could not be done without a larger sample of his construct system. The alternative to some sort of weighting scheme is an assumption of identical importance of all constructs used, i.e., a weight of "one" for each construct. Such an assumption is sure to be erroneous. These two difficulties led to the abandonment of the plan for a checklist and to the eventual adoption of the method of correct matching for comparing protocols.
The Problem as Revised

As a result of the exploratory research the area covered by the problem was narrowed and the hypotheses were revised and made more precise. The original plan was to study the generality of constructs as it might be shown by occurrence of the same or related verbal symbols on several projective techniques. In line with the necessity to limit the problem, it was decided to concentrate most attention on characteristics of constructs elicited by picture materials. The hypotheses have been set up to investigate the generality of constructs—both within a single test (a picture test) and in relation to another test (Kelly's Repertory Test). As an aspect of the problem a study of distinguishability of constructs used by various people was required.

The previous plan of studying constructs about the "personality" or "character" of people was continued. In order to reduce greatly the amount of abstracting that the experimenter probably would be required to do, the new design requested the subjects only to describe the pictured people. The usual practice of having subjects tell stories about the pictures was abandoned. It was expected that omission of the story requirement would lead to a greater amount of descriptive material to work with. Also, forcing the subject to do their describing directly rather than indirectly through the medium of plots would decrease the amount of abstracting re-
quired of the experimenter and therefore reduce the likelihood of error in the protocols. If the experimenter had to abstract subjects' constructs from stories, her own constructs would inevitably become entangled with the supposed constructs of the subjects.

One of the questions studied is whether the constructs elicited by the drawings are heavily influenced by the drawings themselves. That is, are the constructs elicited by the drawings perceptibly characteristic of the particular stimulus material? Suppose a number of subjects describe the personalities or characters of one of the figures in each of several pictures. Are the descriptions so specific to the pictures that judges can match the descriptions with the pictures that elicited them with an accuracy above the chance level? Or are the constructs used in the descriptions so general that the same constructs are used in describing several figures, with the result that judges cannot distinguish the descriptions in terms of the pictures that elicited them?

An hypothesis cannot be rigidly deduced from the psychology of personal constructs. However, certain inferences can be made. If a person's psychological processes evolve toward what he construes to be an optimal anticipation of events (Postulate A), then one would expect some relationship between the pictures and the constructs elicited by each one. The phrase "what he construes to be an optimal anticipation" is the "catch". This would suggest that the degree of relationship
might vary from one subject to another. The "degree of relationship" is also a "catch", for it must be determined by someone, and that person's construction of the pictures will influence his ability to discover relationships between the construct systems of the subjects and the pictures which elicit the constructs. This is one sort of situation covered by the Sociality Corollary (A-4), which reads: "To the extent that one person's construction system subsumes that of another he may play a role in a social process involving the other person." (52:50-52) According to this corollary variation is to be expected in the ability of judges to match subjects' constructs with the pictures that elicited them. If a judge can see subjects as construers he is more likely to be able to match pictures and constructs accurately than if he merely agrees or disagrees with subjects' descriptions of the pictures. For this reason, in order to maximize the chances of discovering whatever relationships exist between pictures and constructs about them, it is important that the judges be given some training in personal construct theory. At least for subjects able to function outside a hospital it seems reasonable to suppose that trained judges will be able to perceive a relationship between constructs and the pictures that elicited them. Such a supposition is also in line with recent normative data about themes in the Thematic Apperception Test. Therefore the first hypothesis is that "There are both per-
ceptible relationships between certain TAT-type drawings and the descriptions elicited by those drawings and perceptible distinctions between descriptions elicited by different drawings." That is, the degree of generality expected in subjects' construct systems is not so great that figures in various pictured situations will be seen in almost the same way.

The first hypothesis is concerned with placing a limitation on the generality of constructs. The other hypotheses are concerned with the range of generality (some lines along which generality may occur) and the personal quality of constructs. These variables have been separated insofar as it appeared feasible, but, in part, they have been studied together.

As a person meets new events, he can see them only in terms of his constructs of past events. His construction of a new event shows continuity with his constructions of previous events. The question is, in what kinds of events does he see continuities, or what sorts of generalities do his constructs show. Corollary B-2, the Organization Corollary, states that "A person's tolerance of incompatibility is limited by the permeability and definition of certain superordinate aspects of his system." (53:69-70) Although this corollary speaks in terms of tolerance of incompatibility of constructs, it is safe to infer that compatible aspects, and
so the generality of constructs, also occur in relation to superordinate aspects of the person's construct system.

In our society the public constructs of age and sex are given much emphasis. Therefore, might these constructs be used as superordinate ones in individual's personal construct systems? If so, then we would expect a limitation on the variation, within any one person's construct system, of constructs about a particular age or sex group or age-sex group. It might be expected that some constructs used about one group would not be used about another group. That is, the range of generality of some constructs might be such as to include chiefly people of one sex or people of a relatively limited age group. If this is the case, the judges, when given descriptions by a single subject about a number of different people, should be able to divide the descriptions into groups according to their reference to certain age or sex groups. In the matching task described above, a complicating factor is that many other constructs besides age & sex are stressed in our culture. Insofar as these other public constructs are superordinate constructs in an individual's personal construct system, the relationship between age and sex and so-called personality constructs would be obscured. This obscuring would tend to reduce the accuracy of judges' matching.

Whether or not age and sex are important superordinate constructs used in this way by many people cannot be deduced directly from personal construct theory. But it is at least
consistent with the theory to hypothesize that one line of generality of constructs is age-sex grouping. More formally stated, the second hypothesis reads "There are perceptible relationships between the ways in which any one subject describes several figures commonly said to belong to the same age-sex group and perceptible distinctions between the ways in which he describes figures belonging to different age-sex groups."

In the third hypothesis both the personal quality and the generality of constructs are examined. The question is asked whether constructs elicited by the picture material are both perceptibly characteristic of individual subjects and also show some generality within the individual construct system (hypothesis 3a). If constructs are not perceptibly personal in use, there is no need for determining generality of constructs for individuals. Determination of range of generality for groups would be sufficient. In accordance with Corollary A-2, the Individuality Corollary, that "Persons differ in their constructions of experience" (40-42 manuscript), it is expected that there will be differences in the way that various subjects describe the same pictured figures. (The reader is referred, also, to the comparison of intra-individual and inter-individual consistency in chapter II.) Research has shown repeatedly that individuals do differ in their reactions to stimulus materials. Therefore, if the third hypothesis were intended to test only that fact, the work required would be
largely a waste of time. Two other factors are involved which make a test of the hypothesis more worthwhile. One of the factors is involved in a breakdown of the third hypothesis into sub-hypotheses, while the other factor is the relationship of the third hypothesis to the fourth hypothesis. The relationship of the third and fourth hypotheses will be discussed later after the fourth hypothesis has been presented.

Parts of the third hypothesis are set up to test simultaneously some possible limits to the personal quality and generality of constructs. As a test of hypothesis 3b suppose that each subject's descriptions of all the figures in a single age-sex group are lifted from his total protocol, which is composed of descriptions of figures belonging to several age-sex groups. Suppose, further, that each subject's set of descriptions of figures within the one age-sex group is divided into two parts, and that the two sub-sets of descriptions by several subjects are placed in a pile together. The question is whether judges can differentiate between the descriptions given by different subjects and at the same time see commonality in the descriptions by any one subject well enough to put together the descriptions given by each subject. It is hypothesized that they will be able to do this with better than chance accuracy. As a test of hypothesis 3c suppose that each subject's descriptions of all the figures in a single age-sex group are lifted from his total protocol, and then his descriptions of all the figures in a second age-sex group are
lifted from the remaining part of the protocol. Suppose, also, that the two sets of descriptions by each subject are kept separate from each other and that the two sets by several subjects are placed in a pile together. The question is again whether the judges can differentiate between the descriptions given by different subjects and at the same time see commonality in the descriptions by any one subject well enough to put together the two sets of descriptions given by each subject. It is hypothesized that they will be able to do this with better than chance accuracy.

Both hypothesis 3b and hypothesis 3c involve the general hypothesis (3a) that individuals differ in their construction of events and that the constructs used by any one individual show some generality. However, the test of hypothesis 3a involves the use of heterogeneous material (heterogeneous in terms of the age and sex of the figures described). It is possible that the judges can differentiate between the protocols of various subjects and see relationships between parts of protocols by individual subjects while making use of only part of the heterogeneous material that is available. In testing hypotheses 3b and 3c, both the kind and amount of material available are limited. Hypothesis 3b asks whether constructs used to describe figures within a single age-sex group are both personal and general; the several age-sex groups are tested separately. Positive results could be obtained for hypothesis 3b if the personal and general qualities are limited
within single age-sex groups or if they extend across age-sex lines. Hypothesis 3c is intended to differentiate between these two possibilities.

In the fourth hypothesis the investigation of generality of constructs is extended beyond the one test. The test of generality as shown by two different samples of verbal behavior is the part of this study most directly related to the original experimental plan. As has been stated previously, an attempt to discover a person's constructs through tests is useless unless the constructs are general enough to appear in other situations. The experimenter would have preferred to compare constructs from a test situation with those from a non-test situation, but that was not feasible. Therefore it was decided to compare constructs elicited in two different types of tests, i.e., a picture test and a task requiring sorting real people (the Repertory Test).

It is assumed that in the past the subjects have evolved constructs about real people through making predictions and seeing them validated or invalidated. It is expected that when they are called upon to sort the real people, they will fall back upon constructs about those people that were evolved in the past. Likewise, a good many subjects have probably built up constructs about pictured people as a result of seeing movies, reading illustrated stories, seeing advertisements, etc.; when called upon to describe new pictures, they may draw on these constructs to some extent, at least. Any given constructs may
or may not include both real and pictured people as elements. That would depend on whether a subject had had what he would call similar experiences in relation to both real and pictured people. However, his construing of pictured people is against a background of experience with real people, and his construing of real people includes as part of the background his experiences with pictured people. His constructs about real and pictured people may vary somewhat, but, after all, both situations do involve people. If he has a construct about real people, he may find it useful in making predictions about pictured people, and vice versa.

Although variation and inconsistencies are expected within a person's construct system, within PPC they are held to be subordinate to more permeable aspects of the system (Postulate B in 53, pages 59-61). That is, the variation and inconsistencies are held to be subordinate to an open-endedness of most constructs that allows new elements (in this case either pictured or real people) to be added to old constructs. Therefore it is expected that if two protocols from different tests are obtained from each of a number of subjects and the two protocols from several subjects are placed in a pile, judges will be able to pick out which protocols were given by the same subjects. In a test of hypothesis 4a the entire protocols are to be used, and each one is to include descriptions of several figures from each of several different age-sex groups. To test hypothesis 4b the two protocols are
to be broken into parts, with all the figures in any one part belonging to the same age-sex group. The parts of the picture test and the Repertory Test are to be kept separate. Then suppose that a set of descriptions from the picture test and a set from the Repertory Test for each of several subjects are put into a pile, and that all the descriptions in the pile are about figures in a single age-sex group. It is expected that judges will be able to put together the two sets of pictures given by the same subject with an accuracy better than chance. It is not expected, however, that the matching will be perfect, for, after all, different materials are involved and the subjects may have constructs peculiar to those materials. In the tests of hypotheses 4a and 4b, as in all parts of the third hypothesis, the personal aspect of constructs is also involved. Judges are required not only to distinguish similarities of constructs given on two occasions by certain subjects, but they must also differentiate between the construct systems of the various subjects.

The breakdown of test protocols in hypothesis 4b is done for the same reason that the breakdown is done in hypothesis 3b. In working with the whole protocols it is possible for the judges to achieve success in the matching while using only a part of the material, and hypothesis 4b is one check for this. The breakdown in 4b is also made because of the frequent assumption in clinics that a person's descriptions of pictured figures resemble his descriptions or thoughts about
real people of similar sex and age group. An analysis similar to that done in hypothesis 3c was not planned, because an assumption of similarity of description across age-sex lines is less frequent in clinical settings. Also, it was deemed unnecessary, in view of the test, in hypothesis 4b, of the possibility that judges might use only part of the material in their matching of protocols to test hypothesis 4a.

**Experimental Hypotheses**

In the last few pages questions to be examined have been stated, predictions have been made, and the reasons for the predictions have been explained. Now the four main hypotheses and the sub-hypotheses may all be summarized. Such a summary will point up the four experimental variables whose relationship to the descriptions is being studied: particular drawings, age-sex class membership of the figures, personal quality of construction systems, and type of test (here pictures as stimulus material versus names of real people as stimulus material). The four hypotheses read as follows:

1. There are both perceptible relationships between certain TAT-type drawings and the descriptions elicited by those drawings, and perceptible distinctions between descriptions elicited by different drawings.

2. There are both perceptible relationships between the ways in which any one subject describes several pictured figures commonly said to belong to the same age-
sex group and perceptible distinctions between the ways in which he describes figures belonging to different age-sex groups.

3. There are both perceptible relationships between the descriptions of various pictured figures given by any one subject and perceptible distinctions between the descriptions of the same figures given by different subjects.

a. There are perceptible relationships between the descriptions of various pictured figures given by any one subject when those figures are grouped into two sets, each composed of figures representing several age-sex groups, provided that the age-sex groups are the same in the two sets, and there are perceptible distinctions between the descriptions of the same figures given by different subjects.

b. There are perceptible relationships between the descriptions of various pictured figures given by any one subject when those figures are all commonly said to belong to the same age-sex group, and there are perceptible distinctions between the descriptions of the same figures given by different subjects.

c. There are perceptible relationships between the descriptions of various pictured figures given by any
one subject when those figures are of different age-sex groups, and there are perceptible distinctions between the descriptions of the same figures given by different subjects.

4. There are both perceptible relationships between the descriptions of pictured figures and real people given by any one subject, and distinctions between the descriptions of the same figures and people given by different subjects.

a. There are perceptible relationships between the descriptions of pictured figures and real people given by any one subject when the people in both tests belong to several age-sex groups, provided that the age-sex groups used in the two tests are the same, and there are perceptible distinctions between the descriptions of the same figures and similarly related people given by different subjects.

b. There are perceptible relationships between the descriptions of pictured figures and real people given by any one subject when the people in both tests are all commonly said to belong to the same age-sex group, and there are perceptible distinctions between the descriptions of the same figures and similarly related people given by different subjects.
CHAPTER IV

Testing Method and Procedure

Overview of Experimental Method

As a means of obtaining two samples of subjects' constructs two different tests were used. One of the tests - the one on which most of the analysis was done - consists of a series of nine pictures. The subjects were asked to give a list of adjectives or phrases descriptive of the character or personality of each of twelve people in the pictures. After they had finished the descriptions they were asked to record the antonyms of their words and phrases; this was done in order to obtain dimensions which might be roughly comparable to the constructs elicited by the other test. The other test was a version of the Kelly Repertory Test. The subjects were asked to list names of 12 people known to them who fit specified role titles. They were then given three names at a time and asked to sort the people and tell in what way two were alike and yet different from the third person. All the role titles and pictured people fit into four age-sex groups: male peers of the subjects, female peers of the subjects, males older than the subjects, and females older than the subjects. The subjects were all undergraduate men. The protocols were broken down into parts in various ways appro-
appropriate for testing the hypotheses. These parts were typed and then randomly arranged into sets of three or four protocols by the experimenter. The sets were given to two trained judges, who were asked to match the parts that belonged together. In order to test the hypotheses, the accuracy of the obtained matching was compared with the accuracy to be expected by chance.

The Picture Test

Because of interest in the clinical usefulness of the experimental results, the experimenter chose published pictures which are available for use in clinics, hospitals, and schools. Of the various published pictures the sets most appropriate for white, college-age men are the Murray and Symonds series. Neither set is considered ideal for this research, but the clinical advantages from using published material were thought to outweigh the experimental advantages of specially drawn pictures. Preferably the pictures should be ambiguous enough to produce a wide variety of types of descriptions. Previous research (21,25,28) has shown, however, that many of the Murray pictures elicit primarily negatively toned stories. Comments made to the experimenter about the Symonds cards by both psychologists and non-psychologists have indicated a general impression that the Symonds cards are even more heavily negatively toned than are the
Murray cards. These cards were used, then, with a realization that their relatively limited ambiguity might reduce the opportunity for subjects to use the same constructs in responding to the picture test and the Repertory Test.

Within the Murray and Symonds sets of pictures, specific cards and figures were selected to fit into four age-sex groups. It was desired to select groups with which all subjects may be expected to have had considerable experience. All college students may be expected to have had a good deal of experience with people of both their own generation and an older generation. The degree of experience with people of a younger generation is likely to vary more. Also, all college students may be expected to have had experience with both men and women. The four groups selected for representation were males of an older generation, females of an older generation, males of approximately the same generation as the subjects (to be known here as young males), and females of approximately the same generation as the subjects (to be known here as young females). The age boundaries of the four groups were set only roughly. Figures whose assigned age varied between 14 or 15 and 30 to 35 years old were included in the younger groups. Figures whose assigned age varied between about 34 or 35 years and 70 to 75 years were included in the older age groups. The original selection and assignment of figures to groups was made on the basis of the experimenter's
guess about the probable statements by the majority of subjects. Rosenzweig and Fleming (95) have published data on some of the figures in terms of statements made by at least 20% of their subjects. The experimenter's guesses were checked against their findings, and it was noted that in no cases were there contradictions. No contradiction was expected since the experiment's subjective criteria were actually more stringent than the norms of Rosenzweig and Fleming. As a further check, each subject in the pretest of the method was asked, after finishing the rest of the test, to state the age of each figure. The only disagreement with the experimenter's assignment of figures to groups was by one subject (known to have applied recently for psychotherapy), who disagreed about the age group of one of the male figures.

Selection of the total number of figures was determined by two factors. On the one hand, the number should be large enough to permit the elicitation of a range of constructs from each subject. This required that several different figures be included within each age-sex group. On the other hand, the number should be small enough to permit all subjects to finish their descriptions in the allotted testing time. Testing time, in turn, was influenced by the amount of time subjects were required to serve in an experiment (all subjects were students in an elementary psychology class) and by the other tasks that the experimental design required. A prelimi-
nary try-out of the procedure indicated that subjects could be expected to describe twelve figures in the 35 minutes available. In order to equalize the amount of data about each age-sex group, it was decided to use three figures in each of the four age-sex groups.

In line with these general points about the number and type of figures, the following criteria were set up for the selection of specific pictures and figures.

1. As many pictures as possible should be from the Murray set, rather than the Symonds set, both because the former set is used more widely clinically and because the negative tone of the former set appears to be less pronounced.

2. All the pictures should contain at least one human figure.

3. The pictures should emphasize one or more individuals rather than a group as a whole.

4. The figures should all look like real people whom the subject might be likely to meet. That is, grotesque or bizarre figures should not be used.

5. The figures should not be identifiable in terms of clothing or other characteristics as belonging to particular occupational, economic, social, religious, or nationality groups, but all should be commonly identified as belonging to the white race.

6. The pictures should include only figures about which the majority of subjects agree in terms of sex and age group.
7. As many as possible of the pictures should contain both a young male and another figure from one of the other three age-sex groups selected for study.

Six Murray cards and three Symonds cards were selected in accordance with these criteria. They contained three figures in each of the young female, older male, and older female groups, but eight figures in the young male group. Both in order to equalize the amount of data available for each group and to allow completion of the task within the allotted testing time, it was necessary to reduce from eight to three the number of young males described by the subjects. Selection of the three could be made either randomly or in accordance with set criteria. The latter policy was decided upon so that the method of selection of these figures would be consistent with the method used in the selection of the other figures. The selection was made so that the young males were pictured with one figure from each of the other three groups; where there was a choice within this criterion, the figure that four pretest subjects described as nearest to the age of college students was selected.

The pictures selected are listed below, and they are identified by the published numbers. The descriptions of the Murray cards are those published in his test manual (80:19-20). Since the Symonds manual contains no descriptions of the pictures, the descriptions given here have been written by the
experimenter.

1. Murray 5:

"A middle-aged woman is standing on the threshold of a half-opened door looking into a room."

2. Murray 12M

"A young man is lying on a couch with his eyes closed. Leaning over him is the gaunt form of an elderly man, his hand stretched out above the face of the reclining figure."

3. Murray 4:

"A woman is clutching the shoulders of a man whose face and body are averted as if he were trying to pull away from her."

4. Murray 7BM:

"A gray-haired man is looking at a younger man who is sullenly staring into space."

5. Murray 6BM:

"A short elderly woman stands with her back turned to a tall young man. The latter is looking downward with a perplexed expression."

6. Murray 13MF:

"A young man is standing with downcast head buried in his arm. Behind him is the figure of a woman lying in bed."

7. Symonds A7:

An adolescent boy is facing a woman, who is seen from a rear-side view.

8. Symonds A10:

A young girl and boy are looking at a crystal ball, while an older, costumed man holds his hands over the ball.
9. Symonds A4:

An adolescent boy and an older man are facing each other while the man holds out some money.

The twelve figures selected from these nine pictures are listed below in terms of the four age-sex groups. In the lists the words "left" and "right" indicate the figure on the left or the figure on the right when there are two figures on a card. The word "middle" is used to indicate the middle figure in the one picture in which there are three figures on a card. One picture contains only one figure. The number and letter code indicates the published method of identifying the pictures.

<table>
<thead>
<tr>
<th>Young Males</th>
<th>Young Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>12M - left</td>
<td>4 - right</td>
</tr>
<tr>
<td>6BM - right</td>
<td>A10 - middle</td>
</tr>
<tr>
<td>13MF - right</td>
<td>13MF - left</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Older Males</th>
<th>Older Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>12M - right</td>
<td>5</td>
</tr>
<tr>
<td>7BM - left</td>
<td>A7 - left</td>
</tr>
<tr>
<td>A4 - right</td>
<td>6BM - left</td>
</tr>
</tbody>
</table>

Contrary to the usual practice with these cards, no mention was made of stories. Instead, the subjects were asked to give, orally, a list of three or four adjectives or short phrases which they believed best described the "character" or "personality" of the pictured figures. They were requested to "Tell what kind of person he is in general."
not just what he is like in the situation in the picture. They were asked not to describe physical appearance, occupation, or activity in the picture. They were also asked to indicate, after giving the complete list for each figure, which of the words or phrases described the person's most important characteristic.

After the descriptions were completed, the subjects were asked to look at the pictures quickly again and tell the experimenter "whether any of the figures look similar in the way they're drawn." Notes were made about which figures the subjects considered were drawn much alike and why they thought the figures looked alike. These data were to be used only incidentally in the present experimental design, but the information was obtained with a hope of using them in designing a future study about aspects of the pictures to which the subjects were reacting and about the relationship between their constructs and the "objective" similarities between pictures.

Next the subjects were asked to state the ages and social roles of all the figures, including both those previously described and those not described. The request to state social roles was put in the form of a question to "tell what relationship the people in each picture have to each other. It may be a blood relationship or it may be something else." This information was used to determine whether individual sub-
jects had perceived the figures in the most common age-sex groups. Since part of the experiment involved a study of constructions of people in the four groups, it was necessary to discard subjects who did not perceive figures in the usual groups. In order to have complete protocols for all subjects used, the entire record was discarded for any subject who identified even one figure differently from the norm.

Finally the subjects were given their own lists of descriptive adjectives and phrases and were asked to "write what you consider to be the opposites of all the words in your lists. Write whatever you would use in talking; don't worry about what the dictionary would say." If a subject indicated that he might use two different antonyms, depending upon the situation he was talking about, he was free to look back at the figure that had elicited his original description. Subjects were permitted to refer back to the pictures if they asked to do so, but they were not instructed to do this. Antonyms were requested in order to obtain dimensions with two end-points instead of the single end points or verbal symbols given when subjects merely describe one figure at a time. Dimensions were desired so that the data from the picture test would be comparable, as far as possible, to that elicited by the Repertory Test.
The Repertory Test

The Kelly Repertory Test is a sorting test in which people that the subject has known are grouped and described. This test is a flexible one in that both the method of administration and the kinds of people to be described can be varied. In the test as used here subjects are given a list of role titles and are instructed to write on another sheet of paper names of people they know who fit the titles. Then the title list is taken away. Subjects are given three names at a time and are asked to sort the people and to tell in what way two of the people are alike and yet different from the third person.

Two forms of administration of the test were tried out during the free research period. Nine subjects were tested with the Minimum Context Form (described above), and six were tested with the Full Context Form (in which subjects are given all the names at once, instead of in sets of three, and are asked to place the people in as many groups as they wish and then state the basis of their grouping. The Minimum Context Form was selected for use in the main experiment because: (1) Inspection of the results indicated that, as a whole, subjects who took the Minimum Context Form tended to give a larger number of constructs than those who took the Full Context Form. Also, the number of constructs given varied less widely among those who took the Minimum Context
Form than among those who took the Full Context Form. (2) Subjects who took the Minimum Context Form tended to give a wider variety of constructs. Some of those who took the Full Context Form used a single construct for each sort, and then grouped the names into classes as on a rating scale. (3) The time required for the Full Context Form varied so widely (from five to twenty-five minutes for one sort) as to cause practical difficulties during an experimental session whose length was limited by the time the subjects were required to serve in an experiment.

Exploratory work was done with three different, but partially overlapping, lists of titles and sorts. The original test (the one described by Kelly in *The Psychological Construction of Life*) contained 24 titles chosen from present and high school acquaintances, and it had 32 sorts. During the free research period nine subjects were given all or part of this test; they were given part of the test only when they were not able to complete the entire test during the examination period. It was clear that the lists were too long for the time available for testing subjects in this study. Most of the subjects did not have time to complete the sorts. Also, when such a long task was crowded into a short period, some of the subjects appeared to lose interest in the task and to put less thought and effort into
later sorts than into early ones.

Some method was needed for shortening both the list of titles and the list of sorts. A choice could be made between a logical method and an empirical method. In order to determine whether any of the titles were unusually good or unusually poor for eliciting constructs, the mean number of constructs per sort was determined for each title. For instance if a certain title was used in three sorts by each of the nine subjects, that title appeared 27 times altogether. Now suppose that in these 27 appearances 54 constructs were used. Then for this title the mean number of constructs per sort would be 2.0. The obtained means ranged from 1.58 for the title "the person whom you would most like to be of help to" to 2.18 for the title "the most interesting person whom you know personally". For 16 of the 24 titles the average was between 1.71 and 1.97. These results suggest that although the fruitfulness of individual titles might vary from one subject to another, none of the titles used was really much better or much worse than other titles for the group as a whole. Therefore the decision was made to select titles for a revised list on a logical basis rather than on an empirical basis.

Still in the period of free research a revised list of 18 titles and 18 sorts was given to six additional subjects. Titles referring to specific places or situations in which people were known were omitted from this list (some had been
included in the original list), and titles were added to represent people of various age-sex groups who came nearest to being the subjects' ideals. This list appeared to be satisfactory in terms of time required and elicitation of a variety of constructs. However, the list was discarded and a new one was constructed when the entire experimental design was revised, to provide a more satisfactory test of the new hypothesis.

For the final version of the test a special list of twelve titles was prepared. The titles were all selected with the aim of studying the subjects' constructs about people in the same four age-sex groups that were used in the picture test. In each group there are three titles: one represents the subject himself or a close relative or friend, another represents an ideal person, and the third represents a disliked person. The same types of titles are used in all the age-sex groups so that the subjects' constructs about the four groups will be comparable in terms of the situations that elicited them. In the two earlier lists some titles were inserted even though the constructs elicited by them were not to be used in the proposed experimental design; they were intended to distract the subjects from the experimenter's purpose of obtaining constructs about the four age-sex groups. In the list that was finally used there were no such decoy titles. They were omitted in order to
allow a greater number of comparisons of the figures being studied while the total number of sorts was kept within a limit necessary in terms of the testing period. The change was made with an awareness that under the new procedure the subjects might be more aware of the kinds of comparisons being made and so might use more censorship in stating constructs. Such censorship might reduce the chances of demonstrating generality of constructs used on the Repertory and picture tests.

The specific titles used are listed below in the order in which they were presented to the subjects. The order of presentation was determined randomly.

1. Yourself
2. Your mother. (Or the person who has played the part of a mother in your life.)
3. A male friend or acquaintance, about your own age, who comes nearest to being your ideal.
4. Your present or most recent girl friend. (Or your wife, if you are married).
5. A girl about your own age that you get along well with or would like to know better.
6. An older man that you do not like. (Or one that you did not like in the past.)
7. Your father. (Or the person who has played the part of a father in your life.)
8. A male acquaintance about your own age that you do not like. (Or one that you did not like in the past.)
9. An older woman who is nearest your ideal.
10. An older woman that you do not like. (Or one that you did not like in the past.)

11. An older man who is nearest your ideal.

12. A girl about your own age that you do not like. (Or one that you did not like in the past.)

The subjects were told to "Write the names of the people that fit the titles in the boxes on the Name Sheet. ...This will be your own list. I won't ask to see it, and you may take it with you when you have finished the test today."

The point about secrecy of names was included so that subjects would be less reluctant to describe the people later. The subjects were also instructed not to write down the name of any person more than once so that in the sorting, three different people would always be represented in each set of names. If subjects wished to change names recorded, this was allowed provided the request was made before the directions for sorting the people had been given.

The second and main part of the test consists of the sorting and description of the people, taken three at a time. During the free and fixed design research three different lists of sorts (i.e., lists of sets of three names) were used. Each time that the list of titles was changed, it was necessary to prepare a new list of sorts. The original list contained 32 sorts, and this was found to be too many for most subjects to complete in the available time.
The second list contained 18 sorts; it was possible for most subjects to complete this number. Whereas the grouping of names in the first list was arbitrary, the grouping of names in the second list was done according to certain rules of thumb. The groupings were made so as to force the subjects to compare their girl friends with their parents and with other liked people, their parents with their ideal figures of the same sex, themselves with their parents and with their peer ideals, and various ideal figures with each other.

The third list of sorts (the one used in the fixed design experiment) was prepared in terms of specifications that were still more exacting than those used in the preparation of the second list. It was desired to obtain approximately the same amount of material from the Repertory Test and the picture test in order to maximize the chance for overlap of constructs used on the two tests. In the picture test there are three figures in each age-sex group, and the test directions require subjects to give three or four adjectives or phrases describing each figure. If a subject gave all different adjectives and phrases, a maximum of nine to twelve per age-sex group would be possible. In the revision of the Repertory Test, work done with the two earlier forms of the test served as a rough guide. Testing of nine subjects with the first form of the test had indicated
that under that method of administration approximately 1.8 constructs per sort was the most likely number. Frequently, however, the same constructs were repeated on different sorts. As a result, the number of different constructs elicited per sort was considerably less than 1.8. It appeared that use of three names in each age-sex group and the utilization of each name in four sorts would elicit a maximum of about twelve different constructs per group, i.e., a number comparable to the number elicited by the picture test.

If each sort is composed of three names, and if there are twelve names altogether and each name is used four times, a total of sixteen sorts is required. Sixteen sorts were deemed possible, since preliminary testing had indicated that on the average two minutes were required for each sort. The use of each name four times within the sixteen sorts requires that no name be grouped with any other name more than once. Since each sort contains three names, by this method each person is compared with a total of eight other people out of a possible eleven comparisons.

In addition to these requirements about the number and kinds of sorts, certain other criteria were set up for the preparation of the third sort list. It was considered desirable that the same types of comparisons be made in regard to all the age-sex groups. Furthermore, an attempt was made to treat all the individual people as much alike as possible in
terms of comparisons made. Although strict adherence to these aims was not possible, an approximation to them was made. The list of sorts, as finally used, contains the following types of comparisons: (1) Each person is compared with eight other people, two in each age-sex group. (2) Each principal person (subject, girl friend, father, or mother) and each ideal person is compared twice with other principal people, twice with other ideal people, and four times with disliked people. (3) Each disliked person is compared four times with principal people, and four times with ideal people, but not at all with other disliked people. (4) Within each age-sex group there are eight comparisons involving principal people, eight involving ideal people, and eight involving disliked people; each of these sets of eight comparisons is divided equally among the four age-sex groups so that there are two figures from each age-sex group. These comparisons can be understood best by looking at the list of sorts.

In the list below the title numbers refer to the list of titles given on page 118. The title names are abbreviated so that "Y.M." indicates "young male", "Y.F." indicates "young female", "O.M." indicates "older male", and "O.F." indicates "older female". The abbreviation for the group is written before the name of the type of person within the group.
Sort | Title No. | Title Names
--- | --- | ---
1. | 3 9 12 | Y.M. ideal - O.F. ideal - Y.F. ideal
2. | 2 4 8 | mother - girl friend - Y.M. disliked
3. | 1 7 12 | subject - father - Y.F. disliked
4. | 3 4 6 | Y.M. ideal - girl friend - O.M. disliked
5. | 2 11 12 | mother - O.M. ideal - Y.F. disliked
6. | 7 8 9 | father - Y.M. disliked - O.F. ideal
7. | 2 9 10 | mother - O.F. ideal - O.F. disliked
8. | 5 8 11 | Y.F. ideal - Y.M. disliked - O.M. ideal
9. | 4 7 10 | girl friend - father - O.F. disliked
10. | 1 3 8 | subject - Y.M. ideal - Y.M. disliked
11. | 6 7 11 | O.M. disliked - father - O.M. ideal
12. | 4 5 12 | girl friend - Y.F. ideal - Y.F. disliked
13. | 1 2 6 | subject - mother - O.M. disliked
14. | 3 10 11 | Y.M. ideal - O.F. disliked - O.M. ideal
15. | 5 6 9 | Y.F. ideal - O.M. disliked - O.F. ideal
16. | 1 5 10 | subject - Y.F. ideal - O.F. disliked

From the discussion and the list of sorts it can be seen that the goal of similar treatment of all titles was not achieved. The chief exception occurs in relation to the disliked people. The disliked people are not compared with each other at all, but the principal figures and the ideal people are each compared with four disliked people. The probable result of this is to encourage construing under a superordi-
nate construct "like-dislike". But since that construct by itself will not be classed as an acceptable answer, subjects will be forced to use other, though perhaps subordinate, constructs. Although the same types of comparisons were not possible for all the individual titles, the same comparisons were possible for all the age-sex groups. The latter is actually the more important factor, for the hypotheses involve study of generality of constructs along age-sex lines but not along lines of type of people (i.e., the principal figures or ideal people or disliked people).

In regard to the order of sorts the original plan was that no name should appear in two consecutive sorts. It was hoped that spacing the appearances of any one name would reduce whatever tendency subjects might have to repeat a characterization given previously as "the easy way out" of the task; such a reaction would reduce the number of different constructs given, and that reduction, in turn, would reduce the representativeness of the sample of constructs elicited. After several possible orders were tried out, none was found that fit the desired criterion. In the order selected there are two instances in which the same names appear in two consecutive sorts: the ideal young female appears in sorts 15 and 16, and the ideal older female appears in sorts 6 and 7.

Two methods of presentation of the sorts to the subjects were tried out during prelinary experimentation. Both
methods had been used previously by Hunt (48) in research for a master's thesis. One method involved the use of small cards on which the titles were written; subjects wrote names of people who fit the titles directly on the cards. In the sorting they were given three cards at a time and were asked to group the people and state how two were alike and yet different from the third person. This method, of course, allowed the subject to look at the titles as he was sorting, with the result that he might be strongly influenced by the titles rather than primarily by the names. It also required that the experimenter look at the names that the subject had recorded; and it seemed possible that, even though subjects were permitted to use initials rather than names, some subjects might talk more freely if they knew that no one else would see the names. For these reasons Hunt's stencil method was tried during the latter part of the preliminary experimentation. When this method was used, subjects wrote the names in boxes on a single sheet of paper. For each sort a window stencil was placed over the list of names in such a way that three names appeared in the windows. The subjects then gave their descriptions in terms of the numbers that were used to identify the boxes in which the names were written, rather than in terms of the names themselves. This method also prevented the subjects from looking at the titles while doing the sorting.
After a stencil had been placed over the name sheet, the subject was instructed to "... tell me some important way or ways in which two of the people are alike, but different from the third person. Think of ways that describe their character or personality. Don't describe their physical appearance, or occupation, or whether they are men or women, or whether you like them or not..." He was asked to try to give his answers in terms of single words or very brief phrases and to avoid whole sentences or long descriptions. If he gave lengthy answers - and quite a few subjects did so - he was asked either of two types of question, depending upon the content of his answers. If the answer was merely wordy, he was asked to "say that in a shorter way" on the pretext that the experimenter did not have room on the answer sheet for recording long responses. If the answer was in terms of specific episodes, he was asked, "What kind of person does that sort of thing?" The main purpose in the request for short answers was to force the subject himself to answer at a relatively high level of abstraction. Thus it would not be necessary for the experimenter to abstract "personality" constructs from long descriptions of actions in order to obtain data useful in testing the present hypotheses.

After each subject had finished all the sorts, he was given a questionnaire in which he was asked for information
about the people he had listed on the name sheet. Questions were asked about the ages of all the people and the reasons for liking the ideal figures and disliking the disliked figures. Since the hypotheses involve groups divided, in part, according to age, the experimenter wished to examine the range of ages included in the categories "young" and "older". Such an examination will tell something of the subjects' personal constructs about a dimension that is often treated as a public construct. The information will also make possible a comparison of age ranges among the real people and the pictured people described in the other test. Although the information about reasons for selecting certain people was not to be used in testing the present hypotheses, the experimenter wished to obtain it for exploratory purposes. She was interested in the types of reasons given and the relationship between the reasons and the types of constructs elicited by the test.

Subjects

All the subjects were white, undergraduate men enrolled in the introductory course in general psychology. These students are required, as a part of their course work, to participate in two or three experimental sessions during the quarter, and they received two credits toward fulfillment of that requirement for participation in this study. The students had a choice of experiments in which they could par-
participate, and because of this, some selective factors were present. The first eight subjects were obtained near the end of the winter quarter 1952, and they appeared to be largely students who had procrastinated in filling the requirement of experimental participation. The majority of subjects were obtained during the first half of the spring quarter 1952; particularly at the beginning of the quarter students' comments indicated that many had volunteered in order to "get the requirement out of the way". The subjective impression of the experimenter, however, revealed no outstanding group differences in cooperation, effort, or type of response on the tests. Rather, there were individual differences within both groups. Although selective factors were present here, they did not appear to be related directly to the hypotheses to be tested.

Subjects were obtained by the procedure usual in the psychology course. Dittoed schedules were passed around the class sections by the instructors, and students wrote their names on the schedules to indicate that they wished to serve as subjects in this experiment. The schedules contained brief statements about the kind of subjects required and the type of experiment planned. Subject specifications were listed as: "White men. Willing to participate in 2 sessions, one week apart. (Two credits for the 2 sessions.)" The description of the experiment was intended to differentiate this experiment from other experiments in progress at the time (e.g.,
memory experiments), while still remaining vague enough to prevent the subjects from knowing the exact experimental problem. The description was also intended to reassure the subjects about the anonymity of their answers and so encourage a larger number to volunteer. The description read:

"Study of the opinions of college students through 2 tests, each about 50 minutes long. Test answers will be anonymous, and test results will be analyzed in terms of the group of students rather than in terms of individuals. (This is not a memory experiment.)"

When students signed up for the experiment, they were asked to record not only their names but also their addresses. Post card reminders were sent to students whose appointment times were more than two or three days after the date when they signed the schedule. The use of postcard reminders was found to reduce the number of missed appointments to a worthwhile extent. If a student failed to keep his first appointment, no attempt was made to contact him or induce him to keep later appointments; it was thought that such students might be unreliable about keeping later appointments, and protocols could be used only if subjects took both tests. Post cards were also sent to all subjects between the first and second appointments to remind them of their second appointments. A few subjects failed to keep their second appointments despite the post cards, and these subjects were contacted by telephone. Because of the combination of post cards and telephone calls all subjects but
one came for both tests if they came for the first one.

It was necessary to use subjects that were all of one sex or else to use two groups of subjects with the two sexes placed in different groups. This was a necessary procedure because the sex of the subject may influence his construction system; for constructs are held to change with experience, and people of opposite sex may be expected to have had somewhat different experiences. Since the sex of the subject is not an experimental variable in the study, it must be controlled.

The main consideration that led to the decision to use a single group was the possibility of including a greater number of subjects in the one group; the larger size of the group permits more adequate statistical tests of the hypotheses. Men rather than girls were chosen because the pictures available for men appeared to be more suitable for this experiment. The cards for girls and women included a wider age range in the cards for the so-called young female group. Also, some of the cards for girls are more fantastic or grotesque than those for men, and it was desired to use cards with realistic-looking people.

All subjects were students taking their first quarter of elementary psychology, and all were tested before the class took up the topics of personality and projective techniques. Students in the second quarter of the elementary
course were not used because of previous study of these topics. This kind of selection was made so that students would be naive about projective techniques, particularly the picture test, as many are quite interested in the Thematic Apperception Test once they take up the topic. Since no test similar to the Repertory Test is mentioned in the course, they would be less likely to bring into the experiment preconceived notions about that test; however, they might bring in a greater selfconsciousness about what they were saying as they described the people.

Despite the precaution about scheduling the testing periods, an unknown number of the subjects had previously had an unfavorably oriented introduction to the Thematic Apperception Test. During the first week of the quarter some of the instructors had required each student to write a story about one of the pictures (the particular picture varied with the instructor). The instructors had then returned to all the students supposed analyses of their personalities. They gave the same analysis to all students, after writing it in terms both vague enough and laudatory enough for most students to accept it as genuine. When most students appeared to have accepted the analysis, they were told about the trick. The instructors then used this demonstration as a basis for discussing the practices of fortune-tellers, phrenologists, and other charlatans. In the dis-
cussion some instructors differentiated between such use of picture materials and the more cautious use of picture materials by adequately trained people, while some of the instructors omitted such a differentiation. When the picture test was used in this experiment, some of the subjects spontaneously recalled this demonstration. Their apparent reactions to it varied, as did their willingness to accept the experimenter's explanation that the pictures were being used in a different way in this experiment and that she was interested only in opinions about the pictures and did not intend to try to analyze anyone's personality from his responses. Their comments suggested that most of them accepted the difference, apparently on the superficial basis of the fact that a story was requested in the earlier use whereas only descriptions of people were requested in the present study. Spontaneous comments suggested that a larger proportion of the group tested in the spring quarter than of the group tested in the winter quarter recalled the demonstration. The difference might be based on the length of the time interval between the demonstration and participation in this experiment. No systematic record was kept of the subjects who recalled the demonstration because recall was not seen to be a possible factor until after the experiment was under way. The net result of the earlier experience so far as this experiment is concerned, was probably a greater
awareness of revealing personal ideas about the pictures in contrast to giving supposedly objective descriptions. None of the subjects, however, made any comments suggestive of seeing a connection between the two tests.

The approximate number of subjects used was determined partly by the number required for statistical tests of the hypotheses and partly by the availability of subjects willing to serve in two experiments during a quarter in which many other experiments having shorter testing periods were in progress. The number of subjects also had to be evenly divisible by four, since records of four subjects at a time were to be used in testing hypotheses 3 and 4. In the end, 32 subjects were used in the experiment.

In order to obtain 32 subjects, 49 students were tested. Before the testing was begun certain criteria were set up for the acceptability of subjects. Some of these criteria had to do with the type of reaction to the tests, especially the picture test, and there was no way of foretelling which subjects would meet the criteria without actually giving them the tests. One of the students failed to keep his second appointment, and one of the students was unable or unwilling to do any of the sorts on the Repertory Test (he said he had never thought about the characteristics of people). Fifteen of the students tested were not used as subjects because of failure to meet specifications set up for identification of
one or more figures on the picture test. Because age-
sex group identification of people is one of the experi-
mental variables, it was necessary that all subjects
place each figure in the commonly agreed upon group. This
specification was set up for figures not described as well
as for those described. Even though a subject saw all
the figures that he described as belonging to the usual
groups, if he saw other figures as belonging to unusual
groups, his descriptions might be influenced because he
would see the main figures (the ones he described) as par-
ticipating in different situations. Neither the published
normative material nor the clinical experience of the ex-
perimenter had led her to expect such a large proportion of
students to make at least one deviant identification in the
pictures chosen for study. Twelve students made one deviant
identification, and three students made two deviant identi-
fications. Although the proportion of deviant to total re-
sponses was small (10 out of 282 responses to Symonds cards
and 9 out of 517 responses to Murray cards), deviant responses
were given only to certain figures, and most subjects gave
only one deviant response. It is for this reason that it was
necessary to discard one-third of the students tested. Fail-
ure to anticipate the large proportion of discards is based
partly on the fact that the experimenter's thinking was chiefly
in terms of total proportion of misidentifications per
figure rather than number of persons making even one mis-
identification. Most of the published data is also in terms
of figures on the cards rather than subjects. A further
factor is the experimenter's lack of experience with the
Symonds cards and the fact that she could find no published
normative material about the figures on those cards; although
only one-third of the cards used were Symonds cards, half of
the subjects who made misidentifications did so on those
cards.

A description of the 32 subjects used and the 17 students
tested but not used as subjects will be found in Tables I and II.

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristics of Students Tested</td>
</tr>
<tr>
<td>and Used as Subjects (N = 32)</td>
</tr>
<tr>
<td>Chronological Age</td>
</tr>
<tr>
<td>Range</td>
</tr>
<tr>
<td>Median</td>
</tr>
</tbody>
</table>

It can be seen that the two groups resembled each other in
terms of age, college, and year in college. The subjects used
had a median age of 20.44, and ranged from 18 to 27 years old.
The four undergraduate class years were represented about equal-
ly; the median year in college was 2.94. Subject were from five different colleges and one combined program, with the

greatest number (10) from the College of Commerce. The age range of students who were tested but not used as subjects was from 18 to 28 years, and the median age was 20.58. The group contained a somewhat higher proportion of underclassmen than did the other group; the median year in college was 2.50, but the group included members of all the classes from freshmen through the fifth year (one student was a fifth year engineering student). Students were from the same five colleges as those in the other group, but the representation from the various colleges was more even.

All the subjects in this experiment were treated alike except for rotation of the order of administration of the two tests. Control groups in the usual sense were not need-
ed, since there were no "before" and "after" measures with some form of treatment intervening. For testing hypotheses 1, 2, and 3 no control groups are necessary at all. Determination of chance accuracy of matching protocols has been done by statistical procedures with the result that an empirical determination of chance is not needed. For testing hypothesis 4, which involves matching of protocols from two different tests, control groups are optional. Here the variable for which control groups might be used is the time interval between test administrations. Such a procedure would involve two administrations of the picture test to one group and two administrations of the Repertory Test to another group instead of administration of different tests during the two sessions. Repeat administrations of each test might clarify the factor of instability of construct systems. This is important because the degree of stability of construct systems could influence the extent to which generality of constructs would be shown by two tests given at different times. If construct systems should be quite unstable, then, even though constructs had a high degree of generality at any one time, generality might not show up when the second sample of verbal behavior was obtained after a time interval (in this case one week). Comparison of judges' ability to match protocols from two administrations of the same test and from two different tests would aid in understanding the
two variables of stability of construct systems and influence of the situation in which the constructs were elicited. Such control groups would be important chiefly if the attempt to show generality of constructs between two different tests failed. Even though the use of two additional groups of subjects was considered, the idea was abandoned in the final experimental design. It seemed reasonable to expect that in college age subjects there would not be marked changes in construct systems during an interval of only one week. Also, control of recall of previously given responses appeared to present as much difficulty as possible instability of construct systems.

Testing Procedure

All subjects were tested individually, the first eight in the experimenter's office, and the remaining twenty-four subjects in a small examination room. During all the sessions the subject and the experimenter sat opposite each other at a table. Testing was done in two sessions, each scheduled to last approximately 50 minutes. In 28 of the 32 cases there was an interval of one week between sessions. In four cases holidays made it necessary to vary the interval; one subject was given the second test after only five days, one after eight days, and two after ten days. It was desired that the interval between the two
tests should be long enough to minimize direct recall of responses on the first test during the administration of the second. On the other hand, it was also desirable that the interval should be short enough to minimize changes in the construct systems of the individual subjects as a function of intervening events. The interval of one week was chosen as a convenient compromise. In order to reduce self-consciousness and a desire to hide spontaneous responses and to minimize the giving of socially acceptable responses when they were not the spontaneously occurring ones, all subjects were told that their names would not be put on their answer sheets and that the results were to be analyzed in terms of groups rather than individuals. Both tests were administered orally, and the experimenter recorded responses verbatim.

At the beginning of the first session introductions were made. If the subject appeared to want to do so, the experimenter and the subject talked for a few minutes. If the subject did not respond to overtures of small talk, but appeared to want to "get down to business", they went almost immediately into such preliminary details as setting up a second appointment and arranging for the student to receive credit for participation. As an introduction to the tests themselves students were told

"You know that we live in a world in which there are a lot of other people. Whether we get along in the world
is affected by whether we understand the other people in it. Since people are complicated, there can be many ways of looking at them. I am interested in finding out how young men see the people in these tests. Students taking psychology 401 have been selected as representative of normal young adults. In the session today you are to take the first of two tests designed to help us learn how people look to young men. I am not going to analyze the test results in terms of the individual subjects who take the tests. Instead, I want to get the opinions of many students and put their answers all together. Therefore I hope you won't mind being merely an anonymous member of a large group, and I hope you won't mind if we don't put your name on your answer sheets.

Immediately after this introduction the first of the two tests was administered. The order of administration of tests was rotated so that there would be no constant influence of one test on the other. Because some students who were tested were not used as subjects, the rotation of subjects used was only an approximate alternation. Seventeen subjects were given the picture test first and fifteen were given the Repertory Test first.

All subjects were given the same introduction to the second testing session, regardless of the test that they had taken during the first session. In the second introduction they were reminded of the "purpose" of the study and of the anonymity of their responses. They were told

"You may remember that during the first session of this experiment I told you that I am interested in finding out how other people look to normal young men. I said that people see each other in different ways and there are many ways of looking at others because people are pretty complicated. You have already taken one test designed to help us understand the ideas of young men
about other people. Today there is to be a different test. As I told you before, I want to get the opinions of many students and put those opinions all together. I am not going to analyze the answers of individual students. For that reason your answers will remain anonymous in what we do today, just as they were anonymous in what we did last week."

All subjects were given the same directions for the two tests, regardless of the test which they took first. For the picture test the cards were placed in a pile, face down, in front of the subject. The following directions were used:

"This test is in several parts. Here is a pile of nine cards with pictures on them. In a minute will you look at the cards, one at a time, and describe some of the people on the cards, I'll tell you which ones to describe. Tell something about the person's character or personality. Do not describe his physical appearance or occupation or what he is doing in the picture. Tell what kind of person he is in general, not just what he is like in the situation in the picture. There are no right or wrong answers. That is because people look at each other in different ways, and I want to get a cross-section of how students see the people in these pictures. I'll be your secretary (again) and write your answers, but please don't make them too long. Make your descriptions in the form of a list of three or four words that best tell what kind of person each figure is. If you cannot think of single words, you may use very short phrases, but try to think of words. After you have described each figure, will you tell me which one of the words describes the person's most important characteristic. Do you have any questions?"

If a subject objected to the test, saying that one cannot determine character or personality from pictures, the examiner agreed with him but added that everyone forms opinions. None of the subjects objected to the task after this further explanation. Frequently subjects answered with two
words that sounded as if they might be synonyms. At such times the subject was asked whether they "mean the same thing", and his decision determined whether they were recorded as synonyms or as different items. Because some subjects misidentified the sex of a few of the figures, the experimenter was careful not to use the pronouns "he" or "she" in reference to those figures until after each subject had indicated his identification. The order in which the figures were described is listed below. The letters "L", "R", and "M" indicate that the figure described is placed on the left, right, or middle part of the picture.

1. Murray 5  
2. Murray 12M - L  
3. Murray 12M - R  
4. Murray 4 - R  
5. Symonds A7 - L  
6. Murray 7BM - L  
7. Symonds A10 - M  
8. Murray 6BM - L  
9. Murray 6BM - E  
10. Symonds A4 - R  
11. Murray 13MF - L  
12. Murray 13MF - E

Further details about the method of administration may be found in Appendix A.

After the subject had finished describing the twelve figures, he was told

"As you were describing the figures you may have thought that some of the drawings were quite similar to each other. Will you look at the pictures quickly again and tell me whether any of the figures look similar in the way they're drawn. I'm talking now about the way the drawings are made, not about your views on their personali-
ties. Just tell me about the ones that impressed you as you were describing the figures a few minutes ago.

If the subject stated that he considered certain drawings similar to each other, he was asked, "What about the ways the figures are drawn makes them similar to each other? If he stated that he did not consider any of them to be similar to each other, his answer was accepted.

Next the subject was asked about the ages and social roles of the figures.

"Now let's go over the pictures quickly once again. This time will you tell me who the people are and about how old they are. When I say 'who the people are', I mean that you should tell me what relationship the people in each picture have to each other. It may be blood relationship or it may be something else. Tell me this for each person, whether or not you described him a few minutes ago. Also, will you tell me whether you had the idea when you were describing the person or whether you thought of it just now."

Usually there was no question about this part of the test except on Murray card 5, which contains only one figure. Often subjects asked, "What do you want me to say about her?"

They were given an answer such as "What role does she play? What could you call her?" or "If you were going to make up a story with this picture as an illustration, what would she be in the story?"

Finally subjects were asked to give the antonyms of their descriptive words and phrases in order that construct dimensions might be obtained.

"There is one more part to this test. Will you take these two papers with your descriptions on them and
write what you consider to be the opposites of all the words in your lists. Use this green pencil so that I can tell the difference between your descriptions of the people in the pictures and your lists of opposites. Write whatever you would use in talking; don't worry about what the dictionary would say.

After the subject had finished listing the antonyms, he was told that that test session was finished. If this had been his first test, he was reminded of the second appointment. If it had been his second test, he was thanked for his cooperation and was reassured that the experimenter would see that he received two credits toward the course requirement of experimental participation.

The Repertory Test, like the picture test, was given in several parts. First the experimenter placed in front of the subject a list of the role titles and another paper on which was dittoed a column of empty boxes identified by numbers. Then the subject was told

"The test today is in several parts. The first part of the test consists of a list of titles that should suggest to you some people that you know. Write the names of the people that fit the titles in the boxes on the Name Sheet. For each title be sure to write the name in the space that has the same number. You may write the full names or just the first or last names or even just the initials, but whatever you write you should be able to identify the people when you use the list a little later. This will be your own list. I won't ask to see it, and you may take it with you when you have finished the test today. The first title reads 'Yourself'; write on the Name Sheet in box number 1 the way you want to identify yourself during this test." (Pause) "Continue with the rest of the list in the same way. You may find as you go through the list that you will think of someone whose name you have already listed. If that happens, write
down the name of another person that the title suggests to you. Do not repeat names. When you have finished, you should have a list of twelve different names. Do you have any questions?

Any points not clear were explained, and then the subject was allowed time to complete his list. The time required for completing the list varied from five to fifteen minutes. Before the second part of the experiment was begun, the subject was asked to check whether the list contained twelve different names. If he had repeated any name, he was asked to substitute the name of someone else.

Next the list of titles was removed, and the pile of stencils to be used in the sorting task was placed before the subject. The experimenter stated

"In the second part of this test I want you to tell me something about these people, taking three of them at a time. Here is a set of sixteen stencils. In each stencil there are windows that let three names appear when you put the stencil over the Name Sheet." (At this point the experimenter placed the first stencil over the Name Sheet.)

"Will you look at the three names and then tell me some important way or ways in which two of the people are alike, but different from the third person. Think of ways that describe their character or personality. Don't describe their physical appearance, or occupation, or whether they are men or women, or whether you like them or not. I'll be your secretary and write your answers (again), but please don't make them too long. Try to think of single words to describe them; if you can't think of single words, you may use very short phrases, but please don't use whole sentences. Now let's do stencil number 1. You should be able to see the names of the people numbered 3, 9, and 12. Be sure that the numbers on the stencil and on your Name Sheet corre-
spond. Tell me which two of these people are similar in some respect and yet different from the third person. You may put together any two people you wish. Give me just the numbers of the people, not their names.

If the subject both sorted the people and stated the way in which two were alike but different from the third, his answer was recorded. If he merely sorted the people, he was requested, "Now tell me how the two are alike and how the third is different from them." If the subject still did not understand, he was asked to "Pick a way in which two people are alike and the third person is sort of opposite to the others." If a subject gave a long answer in terms of specific incidents, he was asked, "What kind of person would do that?" or "Can you say that in a shorter way? I haven't room to write all that here." If subjects expressed concern about the similarities or differences between their responses on various sorts, they were told that it did not matter whether they gave the same answers or different ones, and that what was important was whether the descriptions fit the people. Most subjects gave one or two characterizations for each sort. An occasional subject gave many responses for a single sort. Such a subject was asked to select the one or two most important characterizations of the people he had just been describing and then limit his descriptions on later sorts to the
most outstanding points. It was necessary to do this in order to prevent accuracy of matching protocols (in the next part of the experiment) on the basis of length of protocol. Further details about the method of administration of the test may be found in Appendix A.

When all the sorts had been completed, the subject was given a questionnaire in which he was asked for information about the people he had listed on the Name Sheet. The experimenter explained

"There is one more part to this test. Will you answer a few questions about the people whose names you have written on your Name Sheet. Just answer the questions; do not put the names of the people on this questionnaire. There is a separate set of questions for each name, and the numbers of the questions correspond to the numbers of the names."

Although the experimenter had written the subject's responses during the sorting, the subject himself wrote his answers to the questionnaire. A copy of the questionnaire may be found in Appendix A.

Most subjects did not ask for any explanation about the purpose of the experiment other than the vague one they were given in the introduction to the testing sessions. However, if they did ask, they were told that they would be given an explanation at the completion of the two tests. If, at the completion of the two tests, they again requested an explanation about the purpose of the study, they were given an explanation that was intended both to give them a possible sense of participation in an understandable and worthwhile pro-
ject and to divert them from making any connection between the two tests in case they should discuss the experiment with anyone else who might serve as a subject later. In an informal discussion the use of tests in clinics and hospitals was brought up and the need for norms derived from normal people was mentioned. They were told that the title test was a new test just being developed and that the picture test was an old test being used in a new way. All the subjects to whom this explanation was given appeared satisfied.

Pretest of Method and Procedure

At the time the pretest was begun the materials, method, and procedure were approximately like the ones used in the fixed design and reported in the preceding part of this chapter. However, some changes were still being made. Eight subjects were used in this part of the experiment.

The same set of pictures was used as in the final design, but the first four subjects described all eighteen figures on them instead of only twelve figures. Their responses were used as guides in the selection of three young male figures from the group of eight available figures. As stated previously, the pictures were selected so that a young male would be pictured once with an older male, once with a young female, and once with an older female figure, and
so that preference would be given to figures which the sub-
jects described as being nearest to the age level of college
students. The selection was made before the figure de-
scriptions (which were written by the subjects) were con-
sulted by the experimenter. The first five subjects in the
pretest group wrote their descriptions and the last three
gave their descriptions orally. It was found that even
though the test directions stressed making a list of charac-
teristics of the pictured people and did not mention writ-
ing a story, some subjects wrote miniature stories; also,
written administration made control of the amount of material
difficult. Only the last four subjects were requested to
complete the list of antonyms.

More drastic changes were made in the Repertory Test
than in the picture test during this stage of the experiment.
The final (third) form of the test was being worked on while
the first three subjects were being tested, and so they were
given the second form of the test (reported in the section
on exploratory research in chapter III). The last five sub-
jects were tested with the final form. A mixture of written
and oral administration was used, in which some subjects re-
ceived only one form of administration and some subjects re-
ceived a combination. It was found that when written adminis-
tration was used, subjects often stated a way in which two
people were alike and then used an entirely separate dimension
for stating the way in which the third person differed from the other two people. Also, some of them did not follow the directions in regard to the type of construct to be used but described the people in terms of physical appearance, occupation, etc. When oral administration was used the experimenter could remind the subjects of the test directions when necessary and could question the subjects about points that were not clear.

By the time the eight subjects had been tested, the method and procedure appeared satisfactory for use in the main experiment. That method is the one described in the major part of this chapter. After this point all subjects tested were included in the chief experimental group, and no more changes were made in the procedure.
CHAPTER V

Matching Method and Procedure

Alternative Treatments of Test Protocols

Generality of personal constructs elicited by two tests may be studied in any one of several ways. Test procedures other than the one used here might have been used for eliciting test responses. For instance, subjects might have been given one test by the method used here, and from the constructs derived from that test a list could have been made up for each subject; then, in responding to the second test, each subject could have been required to check appropriate adjectives from the list rather than to give constructs spontaneously on the second test. Although such a check list will give a more objective score than the method used in the present design, there is a major disadvantage in the use of a check list: even if the subjects are told that they may use other adjectives not on the list, the method is likely to show a higher proportion of overlap of constructs used on the two tests than is the case when constructs are given spontaneously, for it involves both an element of suggestion and a ready path for escape from the use of much effort on the second task.

In addition to ways of handling responses that require
a test method somewhat different from the one used here, there are at least two ways of dealing with the materials obtained by the present test method. One way involves the construction by the experimenter of a check list of descriptive words and phrases derived from one test; the subject would give his responses to the second test spontaneously and then the experimenter would check or count the number of words or phrases in the first test which appeared in the second test. This method gives a score purported to measure overlap of terms or generality of constructs, but it also has some important disadvantages. Because it is not a predictive measure it is subject to influence from the experimenter's bias. Even though the experimenter sets up criteria for determining overlap, there are likely to be many situations where a large element of judgment is required, and where judgment is needed, bias is possible. Also, criteria for determining overlap are difficult to set up in a practical fashion. One possible assumption is that for any given subject each word represents a different construct, and each time any given word is used, the same construct is being used. This, however, does not allow room for the use of two or more verbal labels for a single construct, nor for the use of a single label for different constructs.

Alternative to a check list or counting procedure is
the method of correct matching. Vernon (124) has advocated this method for the investigation of qualitative aspects of personality. It requires that judges, who are given one or more sets of materials, select the materials which belong together according to some predetermined criterion. The accuracy of the judges' matching is then compared with that to be expected in terms of chance or some other factor. In the present experiment, e.g., judges might be given a set of materials consisting of two protocols by each of four subjects along with instructions to select the two protocols that were given by each subject. After they had completed a number of sets, their accuracy on all the sets might be compared with the accuracy to be expected from the operation of chance alone. The method has the disadvantage of being less precise (at least superficially) than a check list or counting procedure. In this experiment, for example, each protocol consists of several parts, and there is the possibility that judges might match materials on the basis of only part of the data in each protocol. It is difficult to know the extent to which this is done, since the judges themselves may not be entirely aware of the basis of their matching. On the other hand, however, the method has certain advantages over the counting method. For one thing, it is a predictive measure and therefore less subject to influence from judges' bias. Judges' beliefs may enter into their pre-
dictions, but the accuracy of the predictions is determined by some outside agency. Another relative advantage is that the question of the identity of constructs used on the two tests need not be handled for every word or phrase appearing in the protocols. If judges have considerable doubt in the case of a certain construct, they can avoid making a decision on the basis of that construct alone and can match the materials on the basis of other constructs. Although this involves an error of omission, it does not involve the repeated errors of commission which must be made when judges are required to express decisions that give an appearance of greater accuracy and higher level of confidence than is present. Furthermore, the matching method allows judges to use constellations or configurations of constructs, whereas they cannot do that when they must check each construct individually. Another relative advantage is that there is no necessity for assigning each construct a specific weight; instead of this, a decision can be made on a more global basis. Although this has the dangers of vagueness, it also avoids the deception of an impression of greater accuracy than is present.

The Method of Correct Matching

Several different methods of matching were considered before the present one was adopted. All the methods in-
involved difficulties in terms of carrying out the procedures, determining the significance of the findings, or interpreting the results, but the design adopted appeared to be the best one. The method used is what Vernon (123, 124) has called a t:it design. It involves setting up two groups of materials, with the same number of items in both groups. Judges match each item from the first group with one item from the second group. Both the number of items in the groups and the number of sets of two groups can be varied. In the present experiment an "item" may consist of an entire protocol or a part of a protocol by one subject or one of the pictures used to elicit a protocol. In this research an important advantage of the t:it design over other matching designs is that it may be used in testing all the hypotheses that have been set up. Some of the other plans considered could be used only in reference to certain of the hypotheses. Also, the determination of the significance of the obtained accuracy of matching is easier than in the case of other matching designs.

Despite the advantages there are also some difficulties. In tests of hypotheses 1 and 2 the number of items that can be included in each group of material is both fixed and small. (The reason for this will be clearer when tests of the specific hypotheses are discussed later.) In tests of hypothesis 1 each group must contain three items and in a test of hypothesis
2 each group must contain four items. For the sake of
greater comparability of results from the various hypo-
theses, the tests of hypotheses 3 and 4 also used groups
of four items. Such small numbers of items limit the
sensitivity of the tests, for the proportion of success-
ful matches to be expected from chance alone is high.
Both Zubin (137) and Chapman (19) have shown that when
two or more judges match two groups of t items, the mean
distribution of chance results is one correct, regardless
of the size of t. When t is small, if the task is easy,
the judges will not be able to demonstrate the limit of
their ability, for good matching will not lead to such high
levels of statistical significance as when t is larger (124).
On the other hand, when t is small, if the task is very
difficult, there will also be a relatively large propor-
tion of chance successes; however, there is not so much
difficulty here as in the case of an easy task, since sta-
tistical procedures will reflect the influence of chance.

The sensitivity of the method can be increased by using
what Vernon has called a t1:2t design. In that type of de-
sign the two groups of material contain an unequal number of
items. Sensitivity is increased through a reduction in the
number of successes to be expected by chance. But in order
to use this design in testing hypothesis 1 it would be ne-
cessary to increase the length of the picture test to twice
its present length in order to obtain more figure descriptions in each age-sex group. In using a t-test design to test hypothesis 2 the picture test would have to be lengthened further so as to obtain descriptions of figures belonging to other age-sex groups. For these reasons the use of that design was impractical. Although a t-test design could be used in testing hypotheses 3 and 4 without an increase in the length of the tests, the design was not used. It was thought that use of a single design for testing all the hypotheses was preferable for the sake of greater comparability of results between hypotheses.

Tests of the Null Form of the Hypotheses

Hypothesis 1:
There are neither perceptible relationships between certain TAT-type drawings and the descriptions elicited by those drawings, nor perceptible distinctions between descriptions elicited by different drawings.

For testing this hypothesis one list* of materials consisted of three numbered drawings of people, all of whom are

---

*The word "list" will be used hereafter to indicate one group of materials used in a single matching task even though the materials, as given to the judges, were not arranged in list form. The word "group" is actually more appropriate, but use of that word is likely to cause confusion because of the use of the word in the phrase "age-sex group". The word "set" will be used to refer to all the materials used in a single matching task.
commonly said to belong to a single age-sex group. The list cannot consist of more than three items because only three drawings are available for three of the age-sex group if the test material is restricted to the Murray and Symonds cards. The second list consisted of three lettered descriptions, one description about each of the figures (drawings) included in List 1. All three descriptions in the list were given by a single subject. The task was to match the descriptions with the drawings that elicited them. Because separate sets of materials were required for all four age-sex groups and for all 32 subjects, a total of 128 sets were involved in testing this hypothesis. A single set of materials included descriptions of figures in only one age-sex group in order to rule out matching cues based on distinctions between constructs about different age or sex groups that had nothing to do with the drawings themselves. Each set included protocols by only one subject in order to control the factor of between-subject differences in constructs. Such controls make possible a test of the distinctiveness of the relationship between the drawings themselves and the descriptions elicited by them. Records were kept of the accuracy of matching of each set of materials by each judge. In determining the significance of the judges' matchings, a separate test was made for each of the four age-sex groups. The type of matching required may be illustrated by this list of materials in a single set:
List 1 (Numbered Materials)

3 Drawings - Young Male (12M - L)
(6BM - R)
(13MF - R)

List 2 (Lettered Materials)

Subject 1 - 3 Descriptions - Young Male (12M - L)
(6BM - R)
(13MF - R)

Hypothesis 2:

There are neither perceptible relationships between the ways in which any one subject describes several pictured figures commonly said to belong to the same age-sex group nor perceptible distinctions between the ways in which he describes figures belonging to different age-sex groups.

For testing this hypothesis one list of materials consisted of four numbered slips of paper, each containing a description of one pictured figure; there was one figure from each of the four age-sex groups, and the particular figures selected were chosen randomly. The other list consisted of four lettered slips of paper, each containing descriptions of two figures from one of the four age-sex groups; the descriptions on the four papers were about different age-sex groups. The lists could consist of only four items because the items were based on age-sex group, and only four age-sex groups were included in the experiment. All the descriptions in the two lists were given by one subject. The task was to match the various descriptions that referred to each of the age-sex groups, so that, e.g., all descriptions about young men were placed together, all that referred to
young women were placed together, etc. There were 32 sets of materials, since each set contained protocols by only one subject. The type of matching required is illustrated by the listing below of materials involved in a single set.

**List 1 (Numbered Materials)**

Subject 1 - 1 Description - Young Male (12M - L)
1 Description - Older Male (7BM - L)
1 Description - Young Female (4 - R)
1 Description - Older Female (5 )

**List 2 (Lettered Materials)**

Subject 1 - 2 Descriptions - Young Male (6BM - R)
2 Descriptions - Older Male (13MF - R)
2 Descriptions - Young Female (A10 - M)
2 Descriptions - Older Female (A7 - L)

Hypothesis 3a:

There are no perceptible relationships between the descriptions of various pictured figures given by any one subject when those figures are grouped into two sets, each composed of figures representing several age-sex groups, even though the age-sex groups are the same in the two sets, and there are no perceptible distinctions between the descriptions of the same figures given by different subjects.

List 1 consisted of four numbered slips of paper, each of which contained descriptions of four pictured figures. The people described on each paper belonged to four different age-sex groups, and all four descriptions on any one paper were given by a single subject. List 2 consisted of four lettered slips of paper, each of which contained descriptions
of eight pictured figures. On each lettered paper there were descriptions of two figures from each of the four age-sex groups, and all the descriptions on any one paper were given by a single subject. The four sets of descriptions in List 1 were taken from the protocols of four different subjects, and the four sets of descriptions in List 2 were taken from the protocols of the same four subjects. The particular figures described in List 1 were different from the figures described in List 2, and the division of figures between the two lists was made randomly. The task was to put together the two sections of protocol given by each of the four subjects. For testing this hypothesis there were eight sets of materials, since each set was composed of the protocols of four subjects. Illustrative lists from one set of materials are shown below.

<table>
<thead>
<tr>
<th>List 1 (Numbered Materials)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subject 1 - 4 Descriptions</strong></td>
</tr>
<tr>
<td>Young Male</td>
</tr>
<tr>
<td>Older Male</td>
</tr>
<tr>
<td>Young Female</td>
</tr>
<tr>
<td>Older Female</td>
</tr>
<tr>
<td><strong>Subject 2 - 4 Descriptions</strong></td>
</tr>
<tr>
<td>Young Male</td>
</tr>
<tr>
<td>Older Male</td>
</tr>
<tr>
<td>Young Female</td>
</tr>
<tr>
<td>Older Female</td>
</tr>
<tr>
<td><strong>Subject 3 - 4 Descriptions</strong></td>
</tr>
<tr>
<td>Young Male</td>
</tr>
<tr>
<td>Older Male</td>
</tr>
<tr>
<td>Young Female</td>
</tr>
<tr>
<td>Older Female</td>
</tr>
</tbody>
</table>
Subject 4 - 4 Descriptions - Young Male (13MF - R)
Older Male (7BM - L)
Young Female (A10 - M)
Older Female (6BM - L)

List 2 (Lettered Materials)

Subject 1 - 8 Descriptions - Young Male (13MF - R)
Older Male (12M - R)
Young Female (A10 - M)
Older Female (6BM - L)

Subject 2 - 8 Descriptions - Young Male (12M - L)
Older Male (7BM - L)
Young Female (4 - R)
Older Female (6BM - L)

Subject 3 - 8 Descriptions - Young Male (6BM - R)
Older Male (12M - R)
Young Female (4 - R)
Older Female (6BM - L)

Subject 4 - 8 Descriptions - Young Male (12M - L)
Older Male (12M - R)
Young Female (4 - R)
Older Female (13MF - L)

Hypothesis 3b:

There are no perceptible relationships between the descriptions of various pictured figures given by any one subject when those figures are all commonly said to belong to the same age-sex group, and there are no perceptible distinctions between the descriptions of the same figures given
by different subjects.

List 1 consisted of four numbered slips of paper, each of which contained a description of one pictured figure; the descriptions were taken from the protocols of four different subjects. List 2 consisted of four lettered slips of paper, each of which contained descriptions of two other pictured figures; the two descriptions on each paper were by one subject, but the descriptions on the different papers were from the protocols of the four subjects represented in List 1. All the figures described in both lists were from a single age-sex group; the figures within each group were divided randomly between the two lists. The task was to match numbered and lettered papers so that descriptions given by the same subjects were put together. For testing this hypothesis 32 sets of materials were required, since eight sets were necessary for each age-sex group and there were four age-sex groups. As an illustration of the type of task, the materials that might be used in one set are listed below.

List 1 (Numbered Materials)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Description</th>
<th>Age</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 Description - Young Male</td>
<td>6BM</td>
<td>R</td>
</tr>
<tr>
<td>2</td>
<td>1 Description - Young Male</td>
<td>12M</td>
<td>L</td>
</tr>
<tr>
<td>3</td>
<td>1 Description - Young Male</td>
<td>6BM</td>
<td>R</td>
</tr>
<tr>
<td>4</td>
<td>1 Description - Young Male</td>
<td>13MF</td>
<td>R</td>
</tr>
</tbody>
</table>
List 2 (Lettered Materials)

<table>
<thead>
<tr>
<th>Subject 1 - 2 Descriptions - Young Male</th>
<th>(12M - L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 2 - 2 Descriptions - Young Male</td>
<td>(13MF - R)</td>
</tr>
<tr>
<td>Subject 3 - 2 Descriptions - Young Male</td>
<td>(12M - L)</td>
</tr>
<tr>
<td>Subject 4 - 2 Descriptions - Young Male</td>
<td>(12M - L)</td>
</tr>
</tbody>
</table>

Hypothesis 3c:

There are no perceptible relationships between the descriptions of various pictured figures given by any one subject when those figures are of different age-sex groups, and there are no perceptible distinctions between the descriptions of the same figures given by different subjects.

List 1 consisted of four numbered slips of paper, each of which contained descriptions of three pictured figures of a single age-sex group. List 2 consisted of four lettered slips of paper, each of which contained descriptions of three pictured figures of a single age-sex group. All the numbered descriptions referred to one age-sex group, and all the lettered descriptions referred to another age-sex group. For example, all the numbered descriptions might be about young males, and all the lettered descriptions might be about older females. The numbered descriptions were given by four different subjects, and the lettered descriptions were given by the same four subjects; all the descriptions on any one slip of paper were given by a single subject. The task was to put together numbered and lettered descriptions given by
the same subjects. Comparisons were made between descriptions of young males and older males, young males and older females, and young males and young females. In testing this hypothesis, 24 sets of materials were used, since three kinds of across-group comparisons were made. Each type of comparison required eight sets of materials, because there were 32 subjects and each set used the protocols of four subjects.

Various types of comparisons could be made in testing this hypothesis. It was decided to compare each of the other groups with young males because young male subjects were giving their constructs about their pictured peers and other people as they saw them. The plan adopted also involved comparison of constructs about groups differing only in sex, differing only in age, and differing in both sex and age. The type of matching required is illustrated by the listing below of materials involved in a single set.

List 1 (Numbered Materials)

<table>
<thead>
<tr>
<th>Subject 1 - 3 Descriptions - Young Male</th>
<th>(12M - L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(6BM - R)</td>
</tr>
<tr>
<td></td>
<td>(13MF - R)</td>
</tr>
<tr>
<td>Subject 2 - 3 Descriptions - Young Male</td>
<td>(12M - L)</td>
</tr>
<tr>
<td></td>
<td>(6BM - R)</td>
</tr>
<tr>
<td></td>
<td>(13MF - R)</td>
</tr>
<tr>
<td>Subject 3 - 3 Descriptions - Young Male</td>
<td>(12M - L)</td>
</tr>
<tr>
<td></td>
<td>(6BM - R)</td>
</tr>
<tr>
<td></td>
<td>(13MF - R)</td>
</tr>
<tr>
<td>Subject 4 - 3 Descriptions - Young Male</td>
<td>(12M - L)</td>
</tr>
<tr>
<td></td>
<td>(6BM - R)</td>
</tr>
<tr>
<td></td>
<td>(13MF - R)</td>
</tr>
</tbody>
</table>
List 2 (Lettered Materials)

Subject 1 - 3 Descriptions - Young Female
(4 - R)
(A100 - M)
(13MF - L)

Subject 2 - 3 Descriptions - Young Female
(4 - R)
(A100 - M)
(13MF - L)

Subject 3 - 3 Descriptions - Young Female
(4 - R)
(A100 - M)
(13MF - L)

Subject 4 - 3 Descriptions - Young Female
(4 - R)
(A100 - M)
(13MF - L)

Hypothesis 4a:

There are no perceptible relationships between the descriptions of pictured figures and real people given by any one subject when the people in both tests belong to several age-sex groups, even though the age-sex groups used in the two tests are the same, and there are no perceptible distinctions between descriptions of the same people given by different subjects.

List 1 consisted of four numbered pieces of paper, each of which contained descriptions of twelve figures. List 2 consisted of four lettered papers, each of which contained descriptions of twelve real people known to the subjects. (Descriptions in List 1 were derived from the picture test, and descriptions in List 2 were from the Repertory Test.) On each paper (either numbered or lettered) there were descriptions about three people in each age-sex group, and the descriptions about members of any one age-sex group were placed together; the order of appearance of the
age-sex groups was the same on both numbered and lettered papers. The numbered descriptions were given by four different subjects, and the lettered descriptions were given by the same four subjects; the descriptions on any one paper were all given by a single subject. The task was to match numbered (picture test) and lettered (Repertory Test) materials so that descriptions given by any one subject were put together. It can be seen that the two lists made use of the entire picture test and Repertory Test protocols of the subjects. If generality of constructs can be shown either in any parts of the two records, or in the entire records, it should show up here. Since each set of materials contained the protocols of four subjects, and since there were 32 subjects altogether, eight sets were required for testing this hypothesis. Below is a sample list of one set of materials required to test the hypothesis.

List 1 - Picture Test (Numbered Materials)

Subject 1 - 12 Descriptions - Young Male

(12M - L)
(6BM - R)
(13MF - R)

Older Male

(12M - R)
(7BM - L)
(A4 - R)

Young Female

(4 - R)
(A10 - M)
(13MF - L)
<table>
<thead>
<tr>
<th>Subject 2 - 12 Descriptions - Young Male</th>
<th>Subject 3 - 12 Descriptions - Young Male</th>
<th>Subject 4 - 12 Descriptions - Young Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Young Male (12M - L) (6BM - R) (13MF - R)</td>
<td>Young Male (12M - L) (6BM - R) (13MF - R)</td>
<td>Young Male (12M - L) (6BM - R) (13MF - R)</td>
</tr>
<tr>
<td>Older Male (12M - R) (7BM - L) (A4 - R)</td>
<td>Older Male (12M - R) (7BM - L) (A4 - R)</td>
<td>Older Male (12M - R) (7BM - L) (A4 - R)</td>
</tr>
<tr>
<td>Young Female (4) (A10 - M) (13MF - L)</td>
<td>Young Female (4) (A10 - M) (13MF - L)</td>
<td>Young Female (4) (A10 - M) (13MF - L)</td>
</tr>
</tbody>
</table>
List 1 - Picture Test (Numbered Materials) cont.

Older Female (5)
(A7 - L)
(6BM - L)

List 2 - Repertory Test (Lettered Materials)

Subject 1 - 12 Descriptions - Young Male

Subject (Ideal) (Disliked)

Older Male (Disliked)
(Father) (Ideal)

Young Female (Girl)
(Know Better) (Disliked)

Older Female (Mother)
(Ideal) (Disliked)

Subject 2 - 12 Descriptions - Young Male

Subject (Ideal) (Disliked)

Older Male (Disliked)
(Father) (Ideal)

Young Female (Girl)
(Know Better) (Disliked)

Older Female (Mother)
(Ideal) (Disliked)

Subject 3 - 12 Descriptions - Young Male

Subject (Ideal) (Disliked)

Older Male (Disliked)
(Father) (Ideal)
Hypothesis 4b:

There are no perceptible relationships between the descriptions of pictured figures and real people given by any one subject when the people in both tests are all commonly said to belong to the same age-sex group, and there are no perceptible distinctions between the descriptions of the same people given by different subjects.

List 1 consisted of four numbered slips of paper, each of which contained descriptions of three pictured figures. List 2 consisted of four lettered slips of paper, each of which contained descriptions of three real people. (Descriptions in List 1 were derived from the picture test, and descriptions in List 2 were from the Repertory Test.) All the people described on both the numbered and lettered papers belonged to the same age-sex
group. The numbered descriptions were given by four different subjects, and the lettered descriptions were given by the same four subjects. All the descriptions on any one paper were given by a single subject. The task was to match the numbered and lettered papers so that the descriptions given by the same subjects were put together. Since protocol material by four subjects was used in each set, there were eight sets for each age-sex group. And since there were four age-sex groups, 32 sets altogether were required for testing this hypothesis. The type of matching involved is illustrated by the listing below of materials from one set.

List 1 - Picture Test (Numbered Materials)

Subject 1 - 3 Descriptions - Young Male
(12M - L)
(6BM - R)
(13MF - R)

Subject 2 - 3 Descriptions - Young Male
(12M - L)
(6BM - R)
(13MF - R)

Subject 3 - 3 Descriptions - Young Male
(12M - L)
(6BM - R)
(13MF - R)

Subject 4 - 3 Descriptions - Young Male
(12M - L)
(6BM - R)
(13MF - R)
List 2 - Repertory Test (Lettered Materials)

Subject 1 - 3 Descriptions - Young Male
  (Subject )
  (Ideal )
  (Disliked )

Subject 2 - 3 Descriptions - Young Male
  (Subject )
  (Ideal )
  (Disliked )

Subject 3 - 3 Descriptions - Young Male
  (Subject )
  (Ideal )
  (Disliked )

Subject 4 - 3 Descriptions - Young Male
  (Subject )
  (Ideal )
  (Disliked )

Preparation of Materials for Matching

Matching two lists of materials may be done on many bases unless careful controls are used during the testing of the subjects and the preparation of the protocols for matching. Some of the factors affecting accuracy of matching are relevant to any experiment using the method of correct matching, and some are relevant particularly to the present experiment. These factors will be discussed in turn, and the controls used will be described.

One of the most important factors affecting the accuracy of matching in any experiment is the heterogeneity of the material in each set. In fact, Vernon (124) characterized it as the most important factor. Obviously, if the various items in any one list of materials are quite simi-
lar to each other, matching will be difficult because there will be insufficient basis for distinguishing between them. On the other hand, if the items in any one list differ considerably among themselves, judges can distinguish between them easily and so can have a firmer basis for their attempts to match these items with those from another list. Or if two items in a list resemble each other closely, but are clearly distinguishable from the other items in the list, the judges may confuse those two items when doing the matching. Such errors may be termed "good errors", and they have a different significance from other errors although the usual matching procedure does not allow for them.

Heterogeneity has often been neglected in the past, but it is coming to receive more attention. Vernon suggested handling it through a study of second choices and "good errors" in matching. However, he concluded from a hypothetical experiment (123) that in most experiments the inclusion of additional choices makes little difference in the final result. Some experimenters have tried to handle the problem by varying the composition of each set of materials according to certain stated criteria. This method has been used by Palmer (83), who had judges try to match Rorschach interpretations with case analyses; he states that he tried to control the heterogeneity of groups so that alternative descriptions within a single
group were neither very different from nor very similar to the correct description. A few experimenters have used random selection of materials for each set as a solution. Allport and Vernon (4:224-242) used this approach in a study of graphological analyses and personality sketches. Secord (106) recommended randomization in an article published after the present research was begun. He pointed out that if the sample is large enough (he suggested 40 or more subjects), the heterogeneity in the sets of materials approaches that found in the total population.

Randomization was the control for the heterogeneity-homogeneity problem in this research. During preparation of materials for testing hypothesis 1, the three pictures in any one age-sex group were assigned numbers on the basis of the order in which the pictures had been presented to the subjects. Then the three descriptions elicited by those pictures and given by any one subject were assigned letters by use of a table of random numbers. In the preparation of materials for testing hypothesis 2, a table of random numbers was used to assign the various figure descriptions by any one subject to two lists, with the requirement that within any one age-sex group one description would be placed in one list and two descriptions would be placed in the other list. A table of random
numbers was used again in assignment of code letters and numbers to the items in the two lists. In the preparation of materials to test each part of hypotheses 3 and 4, subjects were divided into groups of four by use of a table of random numbers. In the parts of hypothesis 3 the several descriptions by each subject were divided into two parts on the basis of a table of random numbers. In the parts of hypotheses 3 and 4, within each age-sex group, the order of appearance of the two or three descriptions was randomized. And finally, in the parts of both hypotheses 3 and 4 the assignment of code letters and numbers to the items of each list was made by use of a table of random numbers.

None of the other factors to be mentioned is so universally important as the heterogeneity of the material. The other factors to be mentioned are important specifically in the present experiment, although some of them are influential in some other types of matching study as well. The amount of material is important for two reasons, and it is one of the more difficult factors to control. If the protocols of various subjects are unequal in length, the judges will have more information for matching some items than others, for the longer the protocols are, the greater is the opportunity for generality of constructs to be shown. Length of protocols can also be used directly
as a cue in matching, if the judges conclude, e.g., that a person who gives a long protocol on one occasion is likely to give a relatively long one on another occasion. The solution to one part of this problem was to set up the tests and instructions in such a way that both the picture test and the Repertory Test as whole would elicit about the same amount of material and so that the various parts within the two tests would also contain about the same amounts of material. The method of doing this has already been discussed in Chapter IV in the section describing the two tests. If a certain length has been selected for the longer test, then equalizing the length of the two protocols maximizes the opportunity for generality of constructs to appear. This is important in tests of hypothesis 4. For the same reason, equal length of material is important in tests of hypotheses 2 and 3, where generality of constructs within and across age-sex group lines is being studied. And in hypothesis 1, where protocols are matched with pictures that elicited them, within limits, matching may be expected to be more accurate with long than with shorter protocols, and so for comparability of results, equal amounts of material about all the age-sex groups is necessary. The aspect of inter-individual differences in length of protocols was handled through the test instructions. On the picture test, subjects were
instructed to give three or four words or phrases de-
scribing each figure; if they gave less, they were re-
minded of the instructions, and if they gave more, they
were asked to limit themselves to the three or four most
important characteristics. On the Repertory Test, all
subjects had to give at least one construct for each
sort; some subjects gave two constructs on many sorts,
and this was accepted, but if a subject gave more than
two constructs per sort, he was told that it was not ne-
cessary to give so many responses and was asked to select
the most important one or two responses he had just given,

Certain other incidental features may be used as
cues if they are not controlled. All protocols were typed
so that handwriting was not available as an artificial
basis of matching. Mistakes in spelling were corrected.
Mistakes in grammar were also corrected; e.g., "don't
care" was changed to "doesn't care" because of application
to a single person. If a subject characteristically used
complete sentences or long phrases, despite the test in-
structions, he was asked to express his idea in shorter
form; if a subject had great difficulty in doing this in
any particular response, the experimenter suggested several
alternatives and asked whether any of them stated his mean-
ing. Word form is a potential cue, and so certain standard
word forms were set up for use when the protocols were pre-
pared for the judges. For instance, some subjects typically used the adjective and some the verb form of words; when a subject used a verb, the word was changed to the adjectival form if that was available. Subjects also varied in regard to use of suffixes, and certain standard ones were set up for use with given roots. For instance, "unloyal" was changed to the more frequent "disloyal".

Words and phrases which would specifically identify a figure were also eliminated. Most of these were eliminated during testing by the instructions, with the result that few responses were given which identified a person's physical characteristics, posture, dress, or activity in a picture. When subjects characterized people as "motherly" and "fatherly", they were asked to give synonyms, and only the synonyms were used in the protocols that were given to the judges. When protocols were typed, such words as "his", "himself", "her", and "herself" were omitted, and, if necessary, the words "own" or "self" were substituted. Although it was realized that certain unusual responses might serve as matching cues, they were not changed because of the danger of altering the sense of the subject's intention. Examples are the use of "aggressive-retrogressive" and "independent--undependent" as descriptions and antonyms.
Judges

Ability of the judges is an important factor in any experiment in which judges are used. Vernon (124) has discussed judges' ability as it relates to matching experiments, and the most important aspect which he found was the practice and training which the judges received. He states that those factors have been shown to be important in matching experiments done by Binet, Powers, and Cantril and Rand. He states also that he has found no significant relationship between the sex of the judges and their matching ability or between tests and ratings for sociality and matching ability. The average correlation of ten short matching tests with intelligence was only .20.

In the present experiment the major criterion for the selection of judges was their past training. Originally the plan was to use several advanced graduate students in clinical psychology as judges, and it was thought that some previous familiarity with personal construct theory would be desirable, or even necessary. In preliminary experimentation, however, an advanced student in academic personnel, with no previous knowledge of personal construct theory, was able to match protocols as accurately as any of the clinical psychology students who tried the task. Therefore, it appeared that such a previous theo-
retical orientation was not absolutely necessary. In addition, it was thought that if judges without this point of view could match protocols with a significant degree of accuracy, there would be support for the communicability of the approach. And communicability is important in the usefulness of any theory.

Since it appeared unnecessary for the judges to have a background in the psychology of personal constructs, it was decided to use one judge from another area of psychology as well as one with previous knowledge of personal construct theory. The judge with some previous familiarity with the theory (Judge B) was a woman who had just completed her second year of training in clinical psychology. She was paid for her services. The judge with no previous familiarity with the theory (Judge M) was a man who was about to complete all work for his Ph.D. degree in academic personnel. Although he had had considerable experience working with interview protocols, he had not had any experience with projective test protocols. His services were obtained on a trading basis; i.e., it was arranged that the writer would serve as judge in his research and he would serve as judge in hers.

Since the training of judges is an important variable in ability to match, two two-hour training sessions were
held, and between the two sessions the judges did some practice matching. In the training sessions the aim was two-fold: (1) to communicate to the one judge the relevant aspects of the psychology of personal constructs and to draw the attention of the other judge to those aspects; (2) to orient the judges to the cues available for use in matching. It was hoped that because of this training the judges would approach the task with similar and known sets and that the variability of their matching would be reduced. A copy of the typed instructions for the task as a whole and for the tests of individual hypotheses will be found in Appendix B. The type of training will be discussed in the next section, entitled Matching Procedures.

Training of Judges

As was previously stated, two two-hour sessions were held, and between the two sessions the judges did some practice matching. At the beginning of the first session the judges were given typed instructions in which they were told briefly of the theory that a subject's words are symbols of constructs, that the constructs are dimensional in nature, and that a knowledge of a subject's constructs can give clues as to his expectations of other people and his possible directions of movement. (A copy of the typed instructions will be found in Appendix B). This material
was discussed with them until they appeared to understand it. Then they were given information about the type of materials to be matched. They were told that undergraduate men had been asked to name personality characteristics of both pictured and real people, who varied in age and sex, and that the subjects were also asked to state what they considered to be the opposite of each word or phrase used to describe each person. The materials were described to them in part as follows:

"...In the parts of protocols which you will be asked to match, the various words or phrases used to describe any one figure are listed together, single spaced. The opposite of each word or phrase is written after it in parentheses. The type of description obtained and the antonyms stated are illustrated by the example given below.

dominating (indifferent)
conceited (friendly)
forward (bashful)

In the illustration the stimulus person was described as being dominating, conceited, and forward. The subject considered that 'indifferent' is the antonym of 'dominating', 'friendly' the antonym of 'conceited', and 'bashful' the antonym of 'forward'. 'Dominating' is underlined because that was considered to be the most important characteristic of the particular figure described. In some of the sets of descriptions the most important characteristic will be underlined, and in some of the descriptions no such indication will be given. As you go through the materials you will find that the protocol sheets contain descriptions of from one to 12 figures. Adjectives and phrases describing any one figure will be typed together, single spaced. Double spacing between sets of adjectives and phrases will be used to indicate that more than one stimulus person was described.... On any one protocol sheet all the descriptions will be by a single subject. The actual protocols given by the subjects have been
broken up in various ways. Some of the parts will be identified by letters and some by numbers. You are asked to put together the parts of protocols that belong together, matching lettered with numbered parts...

The judges were told that directions for the matching required to test any specific hypothesis would be found in large envelopes containing the materials to be matched to test that hypothesis. After the description of materials and the general instructions had been read, they were discussed in order to clarify any points that were vague.

A typed explanation of cues to be used and false cues to be avoided had been prepared. This was read and discussed in conjunction with some sample protocols. The detailed list of cues is contained in the "Instructions for Judges" in Appendix B. Possible cues may be summarized as follows: (1) similarity or identity of adjectives or phrases used in two parts of a protocol; (2) similarity of dimensions (i.e., adjectives and their antonyms) used in two parts of a protocol; (3) types of characteristics considered by the subjects to be most important, i.e., characteristics underlined; (4) the judges' own abstractions or concepts about the types of figures described; (5) the level of abstraction used by subjects in describing figures; (6) vagueness versus definitive quality of descriptions; (7) vocabulary level of subjects. Judges were warned not to attempt to use spelling or grammatical
errors as cues, because the experimenter had attempted to correct these, and any remaining were typing errors. They were also warned that when they attempted to match pictures with protocols, they should not consider as the only possible description the set of adjectives they themselves might use about a figure. The judges matched several practice sets and were told the level of accuracy achieved in each set. The bases of their matchings were discussed in terms of both the typed list of cues and some additional cues which they had found useful. Both judges had made use of similarity of descriptions and of dimensions, although Judge B made more use of the dimensions than did Judge M. Both had used the "values" of the subjects (i.e., cues 3 and 4, above) and the vocabulary level of the subjects. Judge M had used level of abstraction more than had Judge B. In addition to the cues listed above, Judge B had used several other cues: (1) "clarity of antonyms" (whether the subject gave as negatives "not _____" or stated the antonyms in more positive form), (2) apparent variety of constructs used about any one person as opposed to different ways of stating what appeared to be closely related ideas, and (3) "analysis of the person versus analysis of the way the person interacts with other people" (this is illustrated, e.g., by the difference in breadth that may be im-
plied in one subject's use of "dominates everybody" and another subject's use of "dominant").

After the first training session the judges, working alone, matched from one to four sets of sample materials for each hypothesis. For hypotheses involving only a little material in each set (e.g., hypothesis 1) the judges matched four trial sets, but for hypotheses involving a good deal of material in each set (e.g., hypothesis 4a), the judges matched only one trial set. The protocols used for the training were obtained from subjects who participated in preliminary aspects of the experiment when the procedure resembled closely the one finally adopted.

A second training session was held after the judges had completed the practice matchings. In that session they were told of their degree of accuracy. The bases of matching and reasons for success or error were discussed in regard to one or more sets testing hypotheses 1, 2, 3b, 4a, and 4b. The selection of sets was largely a function of difficulty experienced by one or both judges. The discussion was chiefly in terms of application of cues put forward in the previous training session. Judge B showed greater accuracy than Judge M. The difference was apparently based largely on the
greater use of dimensions by Judge B than by Judge M, and so further explanations and illustrations of that cue were made to Judge M. It was also brought out that reliance on supposedly common cultural stereotypes about the various age-sex groups was often misleading.

Matching Procedure

The judges were given two kinds of typed instructions about matching procedures. One set of instructions applied to all the matching tasks, and the other set of instructions varied from one hypothesis to another. Copies of all the instructions will be found in Appendix B. Judges were instructed to read the "Introduction", "General Directions", and "Cues for Use in Matching" that have been described in the section entitled "Training of Judges". They were allowed to refer back to those sheets as often as they wished during the matching.

Next the judges were told that they would be given a number of large envelopes and that each large envelope contained all the materials needed for testing one hypothesis or one part of a hypothesis. In each large envelope was a set of directions specific to the given hypothesis, and a number of small envelopes, each of which contained one set of materials to be matched.
Judges were asked to open only one large envelope at a time, to read the directions in it, complete the necessary matching, and return the materials to the envelope before opening the next large envelope. They were also asked to open only one small envelope at a time, to match the lettered and numbered materials in the envelope, record their matching on the answer slip provided in the envelope, and place both materials and answer slip back in the envelope before going on to the next set of materials. They were told never to go back to an envelope previously completed.

Since the order in which matching was done could affect the accuracy of the matching, each judge was given a list containing the prescribed order. The two lists were similar in that both subjects matched materials for the first hypothesis last. This was necessary because in testing hypothesis 1 the protocols are identified for the judges in terms of both age-sex group membership and the subject who gave the descriptions. If pictures and protocols were matched early in the sequence, the judges might remember types of descriptions referring to one or more of the age-sex groups when they did the matching for hypothesis 2. And when they did the matching for hypothesis 3, they might remember which protocols were given by particular subjects.
It was also desirable that the matching for hypothesis 4 be done after that for hypotheses 2 and 3. For hypothesis 4a uses as a unit the entire picture test protocol of each subject, and hypothesis 4b uses large sections of each subject's picture test protocol, while hypotheses 2 and 3 use the protocols as broken into parts and require the judges to put those parts together to form the larger units given to them for hypothesis 4. It was also important that hypothesis 3a be matched after hypotheses 3b and 3c and that hypothesis 4a be matched after hypothesis 4b, because 3a and 4a contain more materials than the other parts of those hypotheses; matching for hypotheses 3a and 4a might be done on the basis of only part of the material, but the judges might then remember the other parts that had to belong together and then use those memories in matching for hypotheses 3b, 3c, and 4b. In addition to these influences peculiar to this experiment, there were other potential sequence effects common to many experiments. There was the possibility that the judges might match later materials more accurately than early ones because of practice from observing and using cues. Or, on the other hand, later materials might be matched less accurately than early ones because of fatigue or boredom. Because of these possible sequence
effects it was desirable to vary the order in which the two judges matched the materials. The orders of matching actually used based, in part, on both classes of requirements. Judge M matched in the order: H2, H3b, H3c, H3a, H4b, H4a, H1. Judge B matched in the order: H3b, H3c, H3a, H4b, H4a, H2, H1. It can be seen that there was partial rotation of order, and that, in the main, the influence of one type of matching on another was considered. It was hoped that any influence from placement of hypothesis 2 after hypothesis 4 for Judge B would be counteracted by the number of subjects and sets of materials involved. Within each hypothesis the order of matching sets was reversed for the two judges. Thus, if there were 32 sets, Judge M might match in the order 1 to 32, and Judge B might match in the order 32 to 1.

Judges were asked to make brief notes about cues that they considered of special importance, particularly if those cues were not listed in the instructions sheet. They were also asked to note briefly any special difficulties that they experienced. Such notes could be of help in the interpretation of the matching results.

The amount of material to be matched during any one period was determined individually by each judge. Both fatigue and time available between other duties entered into the decision. Judge M stated that he planned to work
in periods of about two hours at any one time. The periods used by Judge B varied more.

Judges were not told about the experimental hypotheses until after they had completed all matching. In the directions for testing each hypothesis they were given only general information about the type of material being matched. They were not told, for instance, that hypothesis 4 involved matching of descriptions of real and pictured people. Such specific information was withheld in order to minimize the influence of any bias that the judges might have about the extent to which certain kinds of materials could be matched. Judges were also, of course, not told about the degree of success they were achieving until after all matching was finished.
CHAPTER VI

Statistical Analysis of Matching Data

Tests of the Null Hypotheses

Several different methods of analysis of data from matching experiments were investigated. These can be divided into two groups: methods which attempt to measure the degree of relationship between the two types of materials matched, and methods which attempt to determine the probability that the obtained accuracy of matching could have occurred by chance. All the methods had both advantages and difficulties when used with data of the type obtained from the present design. Therefore several were used in analyzing the data. Each of the methods used and one method not used will be described in turn.

When Allport and Vernon (4:146-151) and Powers (86) investigated personality through matching of character sketches and handwriting, they compared the number of correct matchings obtained with the number to be expected from chance alone. However, in 1934 Chapman (19) criticized their method of determination of chance expectations and pointed out that the "... whole effectiveness (of the method of correct matchings) rests obvious-
ly upon a correct understanding of what is to be ex-
pected from chance matchings . . . " His discussion about
the statistics involved in this method was probably the
first one written for psychologists. His first dis-
cussion was in terms of two lists having equal numbers
of items, but subsequently he discussed the problem in the
general case where the number of items in the two lists
are unequal (18). He presents the derivation of formu-
lae and tables for showing "the probability that a single
arrangement of two series of t elements will give exac-
tly s correct matchings" and "the probability that a
single arrangement of t elements will give s or more
correct matchings". He also derives and presents a table
showing "the probability (p) that the mean number of
correct matchings resulting from n independent trials
will be as great as s or greater, when t = 4n". He pro-
vided a special table of probabilities of various values
of s because he showed that "... the chance distribution
of values of s is not normal, but skewed and leptokurtic;
hence none of the techniques which evaluate a matching
result in terms of a difference or a coefficient and its
P.E., with reference to tables of the normal curve, are
strictly legitimate." This table can be used either when
several judges match a single set of materials or when one
judge matches materials from several subjects. However,
he makes no mention of procedure when several judges match materials from several subjects, and that is the type of method needed in the present study. Secord (107) states that he used Chapman's method in a study involving both multiple judges and multiple subjects, but he does not explain what he did. It is not clear whether he used Chapman's method for multiple judges or multiple subjects, but he gives no justification for use of that method if he did so do. It appears to this writer that the significance of any given obtained results might differ in the two cases when either judges or subjects are multiple and when both are multiple. Although Chapman's method is an exact one, it does not appear applicable to the present study because of the use here of multiple judges and multiple subjects. Also, his table of probabilities does not include the case in which each list contains only three items. In the tests of hypothesis 1 each list could contain only three items.

Soon after Chapman's analysis appeared, Vernon (122,123,124) published his modification of the mean square contingency coefficient for measuring the degree of relationship between the two types of materials that are matched. (This was not the method which he and Allport had used earlier and which Chapman had criticized.) He showed that the use of chi-square with the usual
contingency table led to a measure of the consistency of matchings rather than the accuracy of matchings, and therefore derived his modified contingency coefficient to measure the accuracy of matching. He says that his mean square contingency coefficient is analogous to a correlation coefficient. He gives both formulae and a graph accurate to two decimal places for determining the coefficient and its standard error. The formula for the coefficient is

\[
C = \frac{\sqrt{(S_t - t)^2}}{\sqrt{(t - l) + (S_t - l)^2}}
\]

Where:
- \(C\) = Modified contingency coefficient
- \(t\) = Number of items per list
- \(S\) = Proportion of correct judgments

The same formula and graph are used when either or both judges or subjects are multiple. With this method the maximum possible value of \(C\) is equal to \(1.0\) only when \(t\) is infinite.

Vernon gives a method for determining the standard error of the coefficient which can be used when several judges match one set of materials or when one judge matches several sets of materials. But he says that when several sets of materials are matched by more than one judge, no formula is possible and "... it can only be determined on the basis of observed variations within the available experimental results (except in the hypothetical case"
where there is no inter-correlation between the sets, or between the judges..." (122). He recommends use of the standard error of the average experiment instead of that of the combined experiments. (Here he uses the word "experiment" to refer to the matching of one set of materials by one judge.) For figuring the standard error of the average experiment the formula is

$$\sigma = \frac{E}{\sqrt{N}}$$

Here $N$ is equal to $mt$ (the number of sets of materials times the number of items in each list) when the standard error is being determined for the average set of materials. $N$ is equal to $nt$ (the number of judges times the number of items in each list) when the standard error is being determined for the average judge. $E$ is an arbitrary symbol for a term that can be calculated from the formula

$$E = (t-1)\sqrt{\frac{(St-I)^2 + t}{t [t-1] + (St-I)^2}}$$

More simply, $E$ can be read from Vernon's graph, which is accurate to two decimal places.

Vernon discusses three assumptions involved in the use of this method (122,123). One assumption is that the terms $1/N^2$, $1/N^3$, etc., are "negligible". Both $mt$ and $nt$, he says, should be equal to at least 100. In the present experiment, which is more concerned with the
materials than with the judges, \( mt \) may be considered adequate in tests of most of the hypotheses because it is 96 or higher, but in tests of hypotheses 3a and 4a (in which each list of materials contains four items, and there are eight sets of materials) \( mt \) is only 32. The term \( nt \) is inadequate for testing the reliability of judges, since it is either 6 or 8 in tests of all the hypotheses. Another assumption required by this method is that "... the obtained matching contingency table is a random sample of \( N \) judgments from a population with specified proportions in the different categories". He states that this condition is infringed in all matching experiments because the judgments are not all independent; the final judgment in an ordinary matching experiment is usually reached by a process of elimination. However, he says, ball rolling experiments led him to conclude that the formula for the standard error is adequate despite violation of this theoretical assumption. There are no conditions in the present experiment that are particularly relevant to it. The third assumption is normality of the distribution of values of \( C \). He says that the assumption is not wholly unjustifiable when \( C \) lies between about 0.20 and 0.60. The distribution of values of \( C \) is dependent upon the distribution of \( s \) (the number of correct
matches). The skewness of the distribution of $s$ is not so large as Chapman stated it was, he says, if $s$ is larger than 1 but still much less than perfect; under these circumstances the distribution fits a normal curve fairly closely. Therefore, he holds, the value of $C/\sigma_2$ gives a useful measure of the reliability of a matching result, and the probabilities obtained by reference to a normal distribution table are "approximately of the same order as those given by Chapman..." In the present experiment the values of $C$ obtained for three of the hypotheses lie in the range which Vernon stated to have a near-normal distribution. The values of $C$ obtained in tests of the third hypothesis, however, are near the maximum possible values, and so the distribution cannot be assumed to be normal.

The use of Vernon's method with the present data can be considered to give an approximation to the degree of relationship between the materials and to the significance of that relationship. The method is not entirely appropriate because of violation of the first and third assumptions in tests of part of the hypotheses. The method is also not entirely satisfactory because of the necessity to test the significance of the coefficient in terms of the average set of materials rather than in terms of the combined sets. Because of these difficulties another method of evaluation of the data was sought.
One other possible method involves the use of the critical ratio to test the significance of the difference between the obtained number of correct matchings and the number expected from the operation of chance alone. The use of this method was suggested by Professor D.R. Whitney of the Statistics Laboratory in the Mathematics Department at the Ohio State University. Although the sample involved in testing some of the hypotheses is small, the critical ratio and normal distribution table should be used in place of a t-table because the population parameters can be determined exactly. In the formula for the critical ratio \( CR = \frac{X - \bar{M}}{\sigma} \), \( X \) is the obtained number of correct matches for all the sets, \( \bar{M} \) is the theoretical mean or number expected by chance alone, and \( \sigma \) represents the theoretical variability.

Use of the formula for the number of permutations of \( n \) objects, taken \( r \) at a time \( \binom{n}{r} = \frac{n^!}{(n-r)^!} \) \((23:35)\), shows that when there are three items in each list, as in tests of hypothesis 1, there are six possible ways of paring the items. In any one set the probability of getting none right is \( 2/6 \), the probability of getting three right is \( 1/6 \); it is impossible to get only two pairs right. The expected number right for a single set is one, as shown by the equation

\[
E(X) = 0 \cdot \frac{\%}{4} + 1 \cdot \frac{\%}{4} + 3 \cdot \frac{\%}{4} = 1
\]
The expected variance for one set is also one, according to the formula

$$\sigma^2 = \sum (X - \mu)^2 \cdot \rho(X)$$

$$\sigma^2 = (-1)^2 \cdot (\%_o) + (\%_1) + 2 \cdot (\%_2)$$

If there are 32 sets, then the expected number correct is 32, and the variance is 32. In tests of hypotheses 2, 3, and 4, where there are four items in each list, there are 24 possible ways of pairing the items. In any one set the probability of getting none right is $9/24$, the probability of getting one right is $8/24$, the probability of getting four right is $1/24$. Again, both the expected number right and the expected variance for a single set are one; when there are several sets, the population parameters are equal to the number of sets.

After the critical ratio has been determined, a one-tail test is used in determining the probability that the result could have occurred by chance. The one-tail test (52) is applicable here because the experimental hypotheses predict that matching will be significantly better than chance, not merely that matching will be significantly different from chance.

Use of the critical ratio requires the assumption of normal distribution of matches. This is violated, particularly in the tests where there are four items in each

*This is the formula used by Professor Whitney. The writer has not found it in any texts on statistics that are commonly used by psychologists.*
list, for the theoretical distribution is positively skewed. In a skewed distribution, the one number used to express the standard deviation is a compromise or approximation, and therefore the critical ratio and probability value are also approximations. For this reason the use of the critical ratio is not completely satisfactory in the present experiment. According to Professor Whitney, however, the skew is not great enough here to lead to serious error in the use of the normal curve. As a result of these considerations, this method was used in determining the significance of the obtained results, but it was not relied upon as the only method.

Still another method of handling the data is to use Chi-square to compare the obtained distribution of correct matches in the total group of sets for each hypothesis with the distribution to be expected by chance (23:98-104). In using this method in testing hypothesis 1, for instance, determination is made of the number of sets in which the judges had no successes, the number of sets in which they had one correct pair, and the number of sets in which they had three correct pairs. Then the number of sets of each type to be expected from the operation of chance alone is calculated by the use of the same probabilities that were presented in the discussion of the critical ratio method.
Each of the probabilities is multiplied by the total number of sets involved in the test of the hypothesis. Thus in a test of hypothesis 1, in which there are 128 sets, the expected number of sets in which there are no correct pairs is determined by multiplying 128 by 0.3333, the expected number of sets in which there is one correct pair is found by multiplying 128 by 0.5000, and the expected number of sets in which there are three correct pairs is found by multiplying 128 by 0.1667. Then Chi-square is calculated in the usual way by summing the quantities represented by the square of the deviation between the obtained and expected numbers divided by the expected number. In this use of Chi-square, the only restriction is that the sum of the expected frequencies must be equal to the total number of sets used in testing the hypothesis. Therefore the number of degrees of freedom is found by the formula \(df = (k-1)\), where \(k\) stands for the number of possible types of solution. In tests of hypothesis 1, there are three possible solutions (no pairs right, one pair right, and three pairs right in each set); in tests of the other hypotheses, there are four possible solutions (no pairs right, one pair right, two pairs right, and four pairs right). Thus ordinarily there are two and three degrees of freedom, respectively. In tests of some hypotheses, however, because of small expected fre-
quencies in one or two of the categories, it was necessary to combine categories, and in those cases the degrees of freedom are fewer.

In the use of chi-square two assumptions are required. \(74:197-198\). One assumption is that the distribution of expected frequencies is continuous, and the other is that the distribution of observed frequencies about any given expected frequency follows a normal curve. The first assumption is usually violated, and the second can seldom be checked. Small expected frequencies lead to violation of both assumptions, and so McNemar has recommended that \(E\) always be at least 5, and that preferably \(E\) should be at least 10. In the present experiment, it was possible to obtain expected frequencies of at least 5 by combining categories in most cases. However, in tests of hypotheses 3a and 4a, even when two categories were combined, one \(E\) was only 2.3336, and logical considerations made it impossible to combine more categories so as to obtain a larger \(E\). Obtaining enough subjects to bring the \(E\) above 5 would have meant more than doubling the number of subjects, and that was impossible. In the case of hypothesis 3a the difficulty is not serious, because even by inspection the results can be said to be significantly better than chance, but in the case of hypothesis 4a the use of chi-square is questionable. Although
the method is satisfactory statistically in those cases where categories were combined, such combination is not really desirable logically, because it sometimes reduces the sensitivity of the test. For both these reasons chi-square was not used as the sole test of the hypotheses.

A correction for discontinuity is recommended by McNemar, and therefore its use was considered in the present experiment. However, Professor Whitney stated that it is "optional", and it is also noted that Edwards (23:98–113) sometimes makes the correction and sometimes does not make it. In the end it was not used here because calculations were a little simpler without it.

Reliability or Comparability of the Judges

Rather than retest or split-half reliability, the comparability of the two judges was determined. This is what Logan and Hunt (56) have called "multi-judge reliability". However, their methods involving use of correlations or standard deviations as raw scores with computation of means and sigmas were not followed because of the limited variability possible in the present experiment. Calculating the reliability of the judges by dividing the contingency coefficient by the sigma of the coefficient was used but was not relied upon because of the small number of judges. Such calculation seriously
violates the assumptions involved in the method.

Two principal methods were used for determining comparability of the judges. One method involves the use of critical ratios and a table of normal distribution, and it is comparable to the use of critical ratios in the tests of the hypotheses, as discussed above. For any given hypothesis, the number of times that the two judges agreed in their matching of individual pairs was counted. The number of agreements to be expected by chance alone was also determined. (This is comparable to the agreement of one judge with the correct list of matches, and it has been shown, above, that chance agreement is one pair in each set matched). The significance of the difference between the obtained agreements and the agreements to be expected by chance was tested by use of the critical ratio. As was explained earlier, the critical ratio is applicable here, rather than the t-test, because the population parameters (chance expectancy) are known exactly.

The other method used to test the comparability of the two judges involves comparison of the distributions of their scores by means of chi-square. This method is similar to the use of chi-square to test the null hypotheses. For each judge, for hypothesis 1 the number of sets in which zero, one, or three pairs were matched correctly,
was counted. Then chi-square was found by summing the squares of the differences between the two judges' scores divided by the scores obtained by one of the judges. The scores of Judge B were used as a substitute for expected frequencies, because that was the judge with a background in personal construct theory, and it was desired to see whether the other judge differed significantly from her matching.

Differences between Pictures

In the matching, the particular choice of pictures could be a factor in the generality shown in the protocols, and therefore in the accuracy with which the judges were able to match. There was no way of comparing these pictures with an absolute standard, but the pictures could be compared among themselves. This was done with reference to hypothesis 1, in which the judges attempted to match descriptions with the pictures that elicited those descriptions. For testing that hypothesis, the task was divided into four parts, and each part involved pictures of only one age-sex group. By chance alone the judges might be expected to achieve the same number of correct matches for each of the three pictures in any given age-sex group. The distribution of correct matches actually obtained for the three pictures
was compared with chance expectations by the use of chi-square according to the method explained above. In the same way the distribution of correct matches obtained for the four age-sex groups as wholes was compared with chance expectations; by chance alone, it might be expected that the judges would achieve the same degree of success in regard to all four groups.

**Differences between Subjects**

Differences between subjects in generality of constructs is also a possible factor in the accuracy of the judges' matching. Therefore, two types of comparison between subjects were made. In the cases of hypotheses 1 and 2, in which materials were divided so that only one subject's descriptions would be in any given set, the total number of correctly matched pairs was determined for each subject. The mean and the standard deviation were calculated. For hypothesis 1 the variability was wide enough so that it was possible to use Fisher's $t$ to compute the significance of the difference between subjects in approximately the top and bottom quartiles. For hypothesis 2 the distribution of scores was too small to make such a comparison very worthwhile. The second type of analysis of the data by subjects was correlation between subjects' scores on various
combinations of tests of two hypotheses. This was done to see whether subjects whose protocols were matched accurately in one matching task were matched accurately in another task. The rationale will be discussed more fully in Chapter VII, Experimental Results. Pearson product-moment correlations and significance levels of the correlations were determined for all six possible combinations of the hypotheses.
CHAPTER VII

Experimental Results

Reliability of the Judges

The two principal techniques for determining the comparability of the two judges involve comparison of the degree of agreement of the two judges in terms of total pairs matched and comparison of the patterns of accuracy of the two judges. The two methods are not identical for several reasons. The former involves comparison of the two judges with each other, and the latter involves comparison of each judge with the "key" of correct results. Also, the former utilizes a pattern showing some dispersion.

In terms of gross number of pairs of descriptions on which the two judges made the same matches, the judges agreed with each other to a degree much higher than chance alone would indicate. (See Table III) The percentage range of agreements was from 41.41% for hypothesis 2 to 93.75% for hypothesis 3a. The median percent was 63.28% (for hypothesis 1). In tests of hypotheses 1, 3a, 3b, 3c, and 4b, when their agreements were compared with the number of agreements to be expected by chance by use of critical ratios and a normal distribution table, all the P values
were less than the .0001 level. In a test of hypothesis 2, the P value was at the .0001 level, and in a test of hypothesis 4a, the P value was .0170. All these tests were one-tail tests, because the informal

TABLE III
Agreement of Matching by Two Judges

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Obtained No. Agreements</th>
<th>Expected No. Agreements</th>
<th>Per Cent of Agreement</th>
<th>CR</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-OM *</td>
<td>62</td>
<td>32</td>
<td>64.58</td>
<td>5.299</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>1-OF</td>
<td>57</td>
<td>32</td>
<td>59.38</td>
<td>4.419</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>1-YM</td>
<td>57</td>
<td>32</td>
<td>59.38</td>
<td>4.419</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>1-YF</td>
<td>67</td>
<td>32</td>
<td>59.79</td>
<td>6.187</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>1-total</td>
<td>243</td>
<td>128</td>
<td>63.28</td>
<td>10.164</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>2</td>
<td>53</td>
<td>32</td>
<td>41.41</td>
<td>3.712</td>
<td>.0001</td>
</tr>
<tr>
<td>3a</td>
<td>30</td>
<td>8</td>
<td>93.75</td>
<td>7.779</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>3b</td>
<td>82</td>
<td>32</td>
<td>64.06</td>
<td>8.837</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>3c</td>
<td>85</td>
<td>24</td>
<td>88.54</td>
<td>12.452</td>
<td>&lt; .0001</td>
</tr>
<tr>
<td>4a</td>
<td>14</td>
<td>8</td>
<td>43.75</td>
<td>2.122</td>
<td>.0170</td>
</tr>
<tr>
<td>4b</td>
<td>57</td>
<td>32</td>
<td>44.53</td>
<td>4.419</td>
<td>&lt; .0001</td>
</tr>
</tbody>
</table>

* In this and subsequent tables the four age-sex groups will be indicated as follows: OM - older males, OF - older females, YM - young males, and YF - young females.

Sometimes the judges agreed with each other in matching the correct pairs, and sometimes they agreed with each other when both were wrong, i.e., they agreed in their errors. The percentage range of correct
agreements of the two judges was from 18.75% for hypothesis 2 to 93.75% for hypothesis 3a; the median percent was 46.09% (for hypothesis 1). This measure says nothing about the distribution of agreements. For instance, if the judges should agree on half the total number of matchings made, they could (1) agree on all the matches for half the sets and on none of the matches for the other half of the sets, (2) agree on half the matches for all the sets, or (3) show some intermediate pattern. For these reasons there are limitations to the adequacy of this method of evaluation.

One aspect of the pattern of matching was tested by chi-square. Among the sets for testing any one hypothesis a count was made separately for each judge of the number of sets in which 0, 1, 2, 3, or 4 pairs had been matched correctly. (See Table IV.) Then the number of sets falling in each category was compared for the two judges. In this comparison the matching of Judge B, who was the judge familiar with personal construct theory, was used as a standard or "expected" pattern, and Judge M's pattern was compared with Judge B's in order to determine whether there was any significant difference between the two patterns. It was found that for hypothesis 1 as a whole and hypotheses 2 and 4a there was no significant difference
### TABLE IV
Comparison of Two Judges' Accuracy of Matching Sets of Materials

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Judge M</th>
<th>Judge B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-OM</td>
<td>3</td>
<td>12</td>
<td>-</td>
<td>17</td>
<td>-</td>
<td>3</td>
<td>8</td>
<td>-</td>
<td>21</td>
<td>-</td>
<td>63</td>
<td>71</td>
</tr>
<tr>
<td>1-OF</td>
<td>1</td>
<td>19</td>
<td>-</td>
<td>12</td>
<td>-</td>
<td>4</td>
<td>16</td>
<td>-</td>
<td>12</td>
<td>-</td>
<td>55</td>
<td>52</td>
</tr>
<tr>
<td>1-YM</td>
<td>6</td>
<td>16</td>
<td>-</td>
<td>10</td>
<td>-</td>
<td>6</td>
<td>15</td>
<td>-</td>
<td>11</td>
<td>-</td>
<td>46</td>
<td>48</td>
</tr>
<tr>
<td>1-YF</td>
<td>1</td>
<td>12</td>
<td>-</td>
<td>19</td>
<td>-</td>
<td>1</td>
<td>10</td>
<td>-</td>
<td>21</td>
<td>-</td>
<td>69</td>
<td>73</td>
</tr>
<tr>
<td>1-total</td>
<td>11</td>
<td>59</td>
<td>-</td>
<td>58</td>
<td>-</td>
<td>14</td>
<td>49</td>
<td>-</td>
<td>65</td>
<td>-</td>
<td>233</td>
<td>244</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>9</td>
<td>10</td>
<td>-</td>
<td>3</td>
<td>8</td>
<td>11</td>
<td>10</td>
<td>-</td>
<td>3</td>
<td>41</td>
<td>44</td>
</tr>
<tr>
<td>3a</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>8</td>
<td>30</td>
<td>32</td>
</tr>
<tr>
<td>3b</td>
<td>1</td>
<td>8</td>
<td>9</td>
<td>-</td>
<td>14</td>
<td>0</td>
<td>5</td>
<td>7</td>
<td>-</td>
<td>20</td>
<td>82</td>
<td>99</td>
</tr>
<tr>
<td>3c</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>-</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-</td>
<td>23</td>
<td>87</td>
<td>94</td>
</tr>
<tr>
<td>4a</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>-</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>3</td>
<td>-</td>
<td>1</td>
<td>22</td>
<td>14</td>
</tr>
<tr>
<td>4b</td>
<td>3</td>
<td>14</td>
<td>12</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>7</td>
<td>12</td>
<td>-</td>
<td>10</td>
<td>50</td>
<td>71</td>
</tr>
<tr>
<td>total sets</td>
<td>36</td>
<td>152</td>
<td>37</td>
<td>116</td>
<td>51</td>
<td>39</td>
<td>125</td>
<td>33</td>
<td>130</td>
<td>65</td>
<td>545</td>
<td>598</td>
</tr>
</tbody>
</table>
between the patterns of the two judges (Table V). Hypothesis 1 involves a large enough number of sets to make possible a break-down of judges' agreements in terms of age-sex groups of the pictured figures. When the break-down was done, no significant difference was found between the two judges in matching materials about any of the age-sex groups. There was a significant degree of disagreement between judges' patterns in the matching for hypothesis 4b, and disagreement of borderline significance in the matching for hypothesis 3b. In the case of hypotheses 3a and 3c the patterns could not be compared by chi-square because in both these hypotheses there was a zero frequency in one of the cells in the "expected"

TABLE V

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Chi-Square</th>
<th>df</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-OM</td>
<td>2.216</td>
<td>1</td>
<td>0.146</td>
</tr>
<tr>
<td>1-OF</td>
<td>0.000</td>
<td>1</td>
<td>1.000</td>
</tr>
<tr>
<td>1-YM</td>
<td>0.138</td>
<td>1</td>
<td>0.711</td>
</tr>
<tr>
<td>1-YF</td>
<td>0.554</td>
<td>1</td>
<td>0.468</td>
</tr>
<tr>
<td>1-total</td>
<td>3.438</td>
<td>2</td>
<td>0.184</td>
</tr>
<tr>
<td>2</td>
<td>0.864</td>
<td>2</td>
<td>0.655</td>
</tr>
<tr>
<td>3a</td>
<td>-----</td>
<td></td>
<td>-----</td>
</tr>
<tr>
<td>3b</td>
<td>3.793</td>
<td>1</td>
<td>0.052</td>
</tr>
<tr>
<td>3c</td>
<td>-----</td>
<td></td>
<td>-----</td>
</tr>
<tr>
<td>4a</td>
<td>2.000</td>
<td>1</td>
<td>0.156</td>
</tr>
<tr>
<td>4b</td>
<td>9.800</td>
<td>2</td>
<td>0.009</td>
</tr>
</tbody>
</table>

*Chi-square cannot be calculated for two of the hypotheses because there is a frequency of zero in one cell of the distribution used as the expected frequency. See the text for explanation.
distribution even after two categories were combined; this occurred because judge B had no sets in which there were no pairs right or only one pair right. Inspection of the results for hypothesis 3a suggests, however, that there is no significant difference between the two judges in matching; Judge B matched all 4 pairs correctly in all 8 sets of protocols, while Judge M matched all 4 pairs correctly in 7 sets and matched 2 pairs correctly in the eighth set. In working on hypothesis 3c Judge B matched all four pairs correctly in 23 of the 24 sets and matched two pairs correctly in the twenty-fourth set, while Judge M matched all 4 pairs correctly in 20 sets, 2 pairs correctly in 3 sets, and 1 pair correctly in 1 set; here it may be guessed that the judges would show disagreement in matching that is of borderline statistical significance.

The reliability of the two judges was also determined by the use of the standard error of the contingency coefficient as regards correct matchings only. (Table VI). The sample is obviously not adequate for determining the reliability of matching of judges in general, and so any conclusions must be limited to these two judges. The reliability of these two judges can be considered good in the case of hypotheses 1, 3a, 3b, 3c, and 4a, at a borderline level in the case of hypothesis 4a, and poor
in the case of hypothesis 2. In the case of hypothesis 2, the unreliability is associated with a low contingency coefficient.

TABLE VI
Evaluation of Reliability of Two Judges by Contingency Coefficient

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>C</th>
<th>Sigma of Judges</th>
<th>$C/\sigma_C$</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-OM</td>
<td>.61</td>
<td>.23</td>
<td>2.65</td>
<td>.0040</td>
</tr>
<tr>
<td>1-OF</td>
<td>.43</td>
<td>.35</td>
<td>1.23</td>
<td>.1093</td>
</tr>
<tr>
<td>1-YM</td>
<td>.32</td>
<td>.40</td>
<td>0.80</td>
<td>.2119</td>
</tr>
<tr>
<td>1-YF</td>
<td>.65</td>
<td>.20</td>
<td>3.25</td>
<td>.0006</td>
</tr>
<tr>
<td>1-total</td>
<td>.52</td>
<td>.30</td>
<td>1.73</td>
<td>.0418</td>
</tr>
<tr>
<td>2</td>
<td>.18</td>
<td>.37</td>
<td>0.49</td>
<td>.3121</td>
</tr>
<tr>
<td>3a</td>
<td>.85</td>
<td>.04</td>
<td>21.25</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>3b</td>
<td>.73</td>
<td>.13</td>
<td>5.62</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>3c</td>
<td>.85</td>
<td>.04</td>
<td>21.25</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>4a</td>
<td>.58</td>
<td>.24</td>
<td>2.42</td>
<td>.0078</td>
</tr>
<tr>
<td>4b</td>
<td>.45</td>
<td>.30</td>
<td>1.50</td>
<td>.0668</td>
</tr>
</tbody>
</table>

* A one-tail test was used.

The apparent contradictions in findings on comparability or reliability of two judges are a function of measurement of different aspects of their matchings. In the case of the standard error of the contingency coefficient for hypothesis 2, for example, the unreliability is associated with the difficulty of the task and the consequent low level of accuracy. Although the standard error of the contingency coefficient for hypothesis 2 shows considerable disagreement among the judges as to which pairs were
matched accurately, the use of chi-square to test another aspect of the pattern of accuracy of matching for the same hypothesis shows considerable agreement between the judges: the judges agreed in getting a small number of pairs correct in most of the sets (the mean number correct per judge per set was 1.328 out of a possible 4.000), but they disagreed as to which pairs each got correct. The apparent contradiction between the findings by different methods of analysis is elaborated further by the fact that they agreed with each other 41.4% of the time, but agreed with the correct solution only 18.75% of the time; that is, a little over half their agreements were in error. Agreement in errors (one aspect of consistency of matching by the two judges) is not reflected in the standard error of the contingency coefficient, which reflects agreement in correct matches only, but it is reflected in the critical ratio.

Tests of the Hypotheses

In tests of the null hypotheses the mean accuracy of matching by the two judges was used. Since the major interest is in the materials and their "matchability", the sums of the scores of the two judges cannot be used, for that procedure inflates artificially the amount of data; it can give the impression that twice as many subjects
were tested as were actually utilized. Use of the mean is valid only if the two judges' matchings come from the same population of matchings, and the population is defined by measures of both central tendency and variability. In the comparison of the two judges that is reported above, the percent of agreement may be considered analogous to a measure of central tendency, and the evaluation of patterns of accuracy by chi-square may be considered analogous to a measure of variability. In these terms the two judges appear to be relatively similar in total score (central tendency of the population) in tests of all the hypotheses. The two judges seem to be relatively similar in pattern (variability of the population) in tests of all hypotheses except 3b and 4b. Thus the use of mean scores appears justified in hypotheses 1, 2, 3a, 3c, and 4a. Since use of means does not appear justified in tests of hypotheses 3b and 4b, the scores of the two judges were evaluated separately for those two hypotheses in addition to evaluation in terms of means.

In the analysis of matching results the N used in hypothesis 1 is large enough to allow breakdown of the results in terms of the four age-sex groups. This breakdown was done during all three types of analysis, i.e., in the use of Vernon's modified contingency coefficient, the critical ratio, and chi-square. When the four additional
types of task from hypothesis 1 are added to the seven
types represented by the seven hypotheses or sub-hypotheses,
a total of eleven scores is obtained for each of the three
types of analysis.

For the seven main and sub-hypotheses Vernon's con-
tingency coefficients range between .18 for hypothesis
2 and .85 for hypotheses 3a and 3c, as seen in Table VII.

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Maximum Possible C</th>
<th>Obtained C</th>
<th>Sigma of Material</th>
<th>C/\sigma</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-OM</td>
<td>.82</td>
<td>.61</td>
<td>.06</td>
<td>10.17</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>1-OF</td>
<td>.82</td>
<td>.43</td>
<td>.09</td>
<td>4.78</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>1-YM</td>
<td>.82</td>
<td>.32</td>
<td>.10</td>
<td>3.20</td>
<td>.0007</td>
</tr>
<tr>
<td>1-YF</td>
<td>.82</td>
<td>.65</td>
<td>.05</td>
<td>13.00</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>1-total</td>
<td>.82</td>
<td>.52</td>
<td>.04</td>
<td>13.00</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>2</td>
<td>.87</td>
<td>.18</td>
<td>.09</td>
<td>2.00</td>
<td>.0228</td>
</tr>
<tr>
<td>3a</td>
<td>.87</td>
<td>.85</td>
<td>.02</td>
<td>42.50</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>3b</td>
<td>.87</td>
<td>.73</td>
<td>.03</td>
<td>21.47</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>3c</td>
<td>.87</td>
<td>.85</td>
<td>.01</td>
<td>85.00</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>4a</td>
<td>.87</td>
<td>.58</td>
<td>.12</td>
<td>4.83</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>4b</td>
<td>.87</td>
<td>.45</td>
<td>.08</td>
<td>5.62</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

*A one-tail test was used.

Except for hypothesis 2, all the coefficients are .45 or
above. The maximum possible coefficient is .82 in the
case of hypothesis 1 and .87 in the case of the other
hypotheses. Thus, except for hypothesis 2, a moderate
or high degree of relationship is shown between the materials matched. The standard errors of the coefficients range from .01 for hypothesis 3c to .12 for hypothesis 4a (a hypothesis which involves a smaller number of sets to be matched). Evaluation of the coefficients by use of a table of normal distribution and a one-tail test shows all to be significantly higher than chance alone would lead one to expect. Six of the seven $P$ values are less than .0001, while the $P$ value for hypothesis 2 is .0228. Within hypothesis 1, all four age-sex groups also show a degree of relationship that is considerably higher than chance matching would demonstrate; three of the $P$ values are less than .0001, and the fourth $P$ value, for matching pictures and descriptions about young men, is .0007. So far, the description of the results in terms of contingency coefficients has been based on the mean matching scores of the two judges. Evaluation of the matching for hypotheses 3b and 4b was also made separately for the two judges. In working with material for both hypotheses, Judge B matched more accurately than did Judge M. For hypothesis 3b, $C$ is .67 for Judge M and .77 for Judge B; both of these coefficients are significant at less than the .0001 level. For hypothesis 4b, $C$ is .31 for Judge M,
and this is significant at the .0003 level. For hypothesis 4b, C is .57 for Judge B, and this is significant at less than the .0001 level.

Determination of the significance of matching was also done by use of critical ratios. (See Table VIII) Again in this analysis a one-tail test was used. For all the hypotheses and sub-hypotheses matching was considerably better than chance level. Five of seven P values determined in tests of the main and sub-hypotheses are less than .0001, one is between the .001 and .0001 levels, and one (hypothesis 2) is between the

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>CR</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-OM</td>
<td>6.187</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>1-OF</td>
<td>3.801</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>1-YM</td>
<td>2.652</td>
<td>.0040</td>
</tr>
<tr>
<td>1-YF</td>
<td>6.894</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>1-total</td>
<td>9.767</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>2</td>
<td>1.803</td>
<td>.0359</td>
</tr>
<tr>
<td>3a</td>
<td>8.133</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>3b</td>
<td>10.341</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>3c</td>
<td>13.574</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>4a</td>
<td>3.536</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>4b</td>
<td>5.038</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

* A one-tail test was used.

.05 and .01 levels. Analysis of the age-sex groups in hypothesis 1 shows that in three cases the P value is less than .0001; and in the fourth case, pictures and
descriptions about young men, the analysis shows a P value of .0040. These significance levels are associated with mean proportions of correct matching of the following order: hypothesis 1, (as a whole) .6211; Hypothesis 2, .3320; hypotheses 3a, 3b, and 3c, .9688, .7070, and .9427, respectively; and hypotheses 4a and 4b, .5625 and .4726, respectively. The matching of the two judges for hypotheses 3b and 4b was also analyzed separately. For hypothesis 3b, although Judge B matched more accurately than Judge M (proportions of .7734 and .6406 correct, respectively), the matching of both judges is significantly better than chance at less than the .0001 level. For hypothesis 4b, again Judge B matched more accurately (.5547 correct) than did Judge M (.3906); the P value for Judge B's matching is less than .0001, and the P value for Judge M's matching is .0007.

In the use of chi-square to compare the mean obtained patterns of matching with the patterns to be expected from the operation of chance alone, three categories and two degrees of freedom were used in tests of hypotheses 1, 2, 3b, 3c, and 4b. In tests of hypotheses 3a and 4a the data had to be consolidated into two categories, with one degree of freedom, in order to come as near as possible to fulfilling the requirement for an N of at least 5 in each category. The chi-square tests are
summarized in TABLE IX. For the two hypotheses with 1 degree of freedom, chi-square is 8.095, significant at the .007 level for hypothesis 3a, and 4.301, significant at the .041 level for hypothesis 4a. For the other five hypotheses with 2 degrees of freedom, chi-squares range from 2.471 for hypothesis 2, through

### TABLE IX

**Evaluation of Accuracy of Matching Sets of Descriptions by Chi-Square***

*(Mean of Two Judges’ Scores)*

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Chi-Square</th>
<th>df</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-0M</td>
<td>42.768</td>
<td>2</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>1-OF</td>
<td>14.721</td>
<td>2</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>1-YM</td>
<td>7.059</td>
<td>2</td>
<td>.032</td>
</tr>
<tr>
<td>1-YF</td>
<td>50.641</td>
<td>2</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>1-total</td>
<td>98.482</td>
<td>2</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>2</td>
<td>2.471</td>
<td>2</td>
<td>.122</td>
</tr>
<tr>
<td>3a</td>
<td>8.095</td>
<td>1</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>3b</td>
<td>39.126</td>
<td>2</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>3c</td>
<td>54.915</td>
<td>2</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>4a</td>
<td>4.301</td>
<td>1</td>
<td>.041</td>
</tr>
<tr>
<td>4b</td>
<td>16.800</td>
<td>2</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>

16.800 for hypothesis 4b, to 98.482 for hypothesis 1, All except the results for hypothesis 2 are significant at 

*Available tables contained only three decimal places.*
better than the .001 level.* The $P$ value for hypothesis 2 is .122, a non-significant result. In the breakdown of hypothesis 1 into the four age-sex groups, when 2 degrees of freedom were used, chi-squares significant at less than the .001 level were obtained for the categories older men, older women, and younger women, while the $P$ value for the class younger men is .032. Again the results of hypotheses 3b and 4b were analyzed separately for the two judges. In distribution of matching accuracy, as well as the previously mentioned total accuracy of matching, Judge B deviated more from chance expectations than did Judge M. For hypothesis 3b, however, both judges' patterns show deviation from chance expectations that leads to $P$ values of less than .001. For hypothesis 4b the $P$ value for Judge M's matching is .006, and the $P$ value for Judge B's matching is less than .001.

It may be noted that all the $P$ values are higher (i.e., show lower levels of significance) when chi-square is used than when the contingency coefficient or critical ratio is used. This is a function, at least in part, of

* $P$ values could be determined only to three decimal places, instead of four as in the case of the contingency coefficients and critical ratios, because the available tables of chi-square contained only three decimal places.
the fact that one-tail tests were used in evaluating the contingency coefficients and critical ratios, while the chi-squares used may be said to correspond to two-tail tests, in that they make no allowance for the fact that all deviations in matching were in the expected direction and were not merely deviations from chance matching.

All the null hypotheses except hypothesis 2 are rejected by all three methods of statistical analysis. Hypothesis 2 may be regarded as in a questionable status so far as statistical rejection of the null hypothesis is concerned. Use of the contingency coefficient and critical ratio along with one-tail tests suggests that this null hypothesis, also, should be rejected, although the level of confidence is not so high as in the case of other hypotheses. Use of chi-square, with what corresponds to a two-tail test, suggests that the null hypothesis cannot be rejected; even if the obtained P value is divided by 2 in order to approximate a one-tail test, the new P value of .061 is still of questionable significance. Any conclusion about the statistical significance of hypothesis 2 is considerably less certain than conclusions about the other hypotheses. However, when statistical significance is at a borderline level in this way, the conclusion can appropriately be drawn that for practical, clinical purposes, the relationship between
the two lists of materials is not of much importance.

**Characteristics of Matching in Terms of Pictures**

If the three pictures in each age-sex group are all completely ambiguous, the descriptions elicited by them should not show distinguishable features and the accuracy with which the judges could match the various pictures and descriptions should be the same because matching would be based entirely on chance. In tests of hypothesis 1 the two judges made a total of 134 accurate matches involving the three pictures of older males; theoretically, if the pictures are all completely ambiguous, 44.6667 of the matches should have involved picture 12M-R, 44.6667 should have involved picture 7BM-L, and 44.6667 should have involved picture A4-R. (See TABLE X.) Actually the number of correct matchings was 44, 41, and 49, respectively. Here chi-square is 0.6731, there are 2 degrees of freedom, and P is .715. This indicates that as far as the subjects and judges used in this experiment are concerned, the pictures cannot be considered significantly different from each other in eliciting descriptions of the men. Likewise, no significant differences were found between the pictures in the younger male and younger female groups, where the P values are .638 and .739, respectively. The pictures of older females, however, are not comparable in terms of descriptions elicited, for the distribu-
tion of correct matches differed significantly from chance expectations; \( P \) here is .005. For picture 5 there were 52 correct matches, for picture A7-L there were 25, and for picture 6E2d-L there were 30 correct matches. Thus it appears that descriptions of the woman in picture 5 are frequently distinguishable from the descriptions of the other two pictures as a group, but that the descriptions elicited by pictures 6E2d-L and A7-L are not clearly distinguishable from each other.

A comparison was also made of the pictures and descriptions of the four age-sex groups as units. If the pictures within each group are equally ambiguous, the descriptions within one group should not be more distinguishable than the descriptions within other groups, and theoretically the obtained number of accurate matches should be the same for all four groups. Since the two judges made a total of 477 accurate matches of pictures and descriptions in the test of hypothesis 1, it would be expected, theoretically, that 119.25 of the matches would be of pictures and descriptions of older men, 119.25 would be of older women, 119.25 of younger men, and 119.25 of younger women. Actually, 134 correct matches involved older men, 107 involved older women, 94 involved younger men, and 142 involved younger women. When the obtained distribution was compared with the expected one, chi-square
TABLE X

Comparison of Accuracy of Matching Various Pictures and Their Descriptions (Sum of Matches by Two Judges)

<table>
<thead>
<tr>
<th>Pictures</th>
<th>Obtained No. Correct Matches</th>
<th>Expected No. Correct Matches</th>
<th>Chi-square</th>
<th>df</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Older Males</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12M-R</td>
<td>44</td>
<td>44.6667</td>
<td>0.6731</td>
<td>2</td>
<td>.715</td>
</tr>
<tr>
<td>7BM-L</td>
<td>41</td>
<td>44.6667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4-R</td>
<td>5</td>
<td>44.6667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>134</td>
<td>134.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Older Females</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>52</td>
<td>35.6667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A7-L</td>
<td>25</td>
<td>35.6667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6BM-L</td>
<td>30</td>
<td>35.6667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>107</td>
<td>107.0001</td>
<td>11.5678</td>
<td>2</td>
<td>.005</td>
</tr>
<tr>
<td>Younger Males</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12M-L</td>
<td>27</td>
<td>31.3333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6BM-R</td>
<td>33</td>
<td>31.3333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13MF-R</td>
<td>34</td>
<td>31.3333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>94</td>
<td>93.9999</td>
<td>0.9209</td>
<td>2</td>
<td>.638</td>
</tr>
<tr>
<td>Younger Females</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-R</td>
<td>50</td>
<td>47.3333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AL0-M</td>
<td>49</td>
<td>47.3333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13MF-L</td>
<td>42</td>
<td>47.3333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>142</td>
<td>141.9999</td>
<td>0.6095</td>
<td>2</td>
<td>.739</td>
</tr>
<tr>
<td>All Pictures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Older Males</td>
<td>134</td>
<td>119.2500</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Older Females</td>
<td>107</td>
<td>119.2500</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Younger Males</td>
<td>94</td>
<td>119.2500</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Younger Females</td>
<td>142</td>
<td>119.2500</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>477</td>
<td>477.0000</td>
<td>12.7693</td>
<td>3</td>
<td>.007</td>
</tr>
</tbody>
</table>
was found to be $12.7693$; since there are 3 degrees of freedom, this value of chi-square is significant at the .007 level. Because of this, the four age-sex group pictures cannot be considered equally ambiguous in terms of the descriptions which they elicit. Pictures and descriptions about young men were the hardest to match, while those about young women and older men were relatively easy to match.

**Characteristics of Matching in Terms of Subjects**

Measures of central tendency and variability were obtained for the 32 subjects in relation to the matching done to test hypotheses 1 and 2. In these two tests, only one subject’s descriptions were included in each set of materials to be matched. Therefore, evaluation of the accuracy of matching of any one subject’s descriptions is not confused by the heterogeneity or homogeneity of subjects placed in the set with him but is a function of the heterogeneity or homogeneity of his own descriptions of the various pictured figures. Since the third and fourth hypotheses involve protocols of four subjects in each set, homogeneity or heterogeneity of subjects placed together is a factor in accuracy of matching. For that reason, the results of tests of those two hypotheses were not used in the analysis of similarities and differ-
ences between subjects.

For hypothesis 1, 24 matches were made for each subject, 12 by each judge. The range of pairs correct per subject was from 10 to 20. Chance number correct is 8 pairs. The mean obtained number correct per subject was 14.875, with standard error 0.537, and the standard deviation of the distribution was 3.038. For hypothesis 2, 8 matches were made for each subject, 4 pairs by each judge. The range of pairs correctly matched per subject was from zero to eight. Chance number correct is 2 pairs. The mean number of pairs matched correctly per subject was 2.656, with a standard error of 0.470, and the standard deviation of the distribution was 1.864. These statistics are summarized in Table XI.

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>Possible Correct</th>
<th>Range</th>
<th>Mean</th>
<th>Sigma</th>
<th>S.E. of Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>10-20</td>
<td>14.875</td>
<td>3.038</td>
<td>0.537</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>0-8</td>
<td>2.656</td>
<td>1.864</td>
<td>0.470</td>
</tr>
</tbody>
</table>

*Analysis based on sum of matchings of two judges

The obtained distribution of number of pairs matched correctly for each subject for hypothesis 1 is quite irregu-
lar, but its form suggests that if a large number of subjects were tested and if their descriptions were matched, the distribution would be platykurtic and perhaps even rectangular. For hypothesis 2, on the other hand, the obtained distribution is positively skewed and appears leptokurtic. It seemed possible that in the case of hypothesis 1 there might be statistically significant differences between subjects in the degree to which their protocols could be matched. Therefore the top and bottom quartiles were compared by means of Fisher's t. In the bottom quartile the mean number of pairs matched correctly was 11.111, with standard error of 0.246, while in the top quartile the mean number of pairs matched correctly was 19.000, with standard error of 0.354. For the difference of 7.889, t is 9.227, which is significant at less than the .001 level when there are 15 degrees of freedom. Thus it appears that the difference in "matchability" of the subjects' protocols is unlikely to be due to chance alone. This type of comparison was not made for hypothesis 2 because of the skewed distribution and the fact that matching was highly accurate for only a small number of subjects.

Since non-chance differences were found between subjects in terms of the matching for hypothesis 1, the question came up as to whether there might be some cor-
relation between "matchability" of subjects' descriptions in tests of two or more hypotheses. For each hypothesis the number of pairs of descriptions matched correctly was determined for each subject. Then the subject's scores on one hypothesis were correlated with their scores on another hypothesis. Product-moment correlations were calculated between all possible combinations of hypotheses, and the significance levels of the correlations were determined. (See Table XII). All the correlations were low. No relationship was demonstrated between matchability in tests of hypothesis 1 and 2 or 1 and 4. A low negative, but statistically significant, relationship appears to exist between subjects' scores on hypotheses 2 and 4, while a low positive, but statistically significant, relationship exists between subjects' scores on hypotheses 3 and 4. Low negative, and probably non-significant, correlations were found between subjects' scores on hypotheses 1 and 3 and 2 and 3.

TABLE XII
Product-moment Correlation between Subjects' Scores on Various Hypotheses

<table>
<thead>
<tr>
<th>Hypotheses</th>
<th>r</th>
<th>r/</th>
<th></th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 vs. 2</td>
<td>-.001</td>
<td>.180</td>
<td>.006</td>
<td>.500</td>
</tr>
<tr>
<td>1 vs. 3</td>
<td>-.220</td>
<td>.180</td>
<td>1.222</td>
<td>.112</td>
</tr>
<tr>
<td>1 vs. 4</td>
<td>.057</td>
<td>.180</td>
<td>.317</td>
<td>.374</td>
</tr>
<tr>
<td>2 vs. 3</td>
<td>-.224</td>
<td>.180</td>
<td>1.244</td>
<td>.103</td>
</tr>
<tr>
<td>2 vs. 4</td>
<td>-.379</td>
<td>.180</td>
<td>2.105</td>
<td>.017</td>
</tr>
<tr>
<td>3 vs. 4</td>
<td>.436</td>
<td>.180</td>
<td>2.422</td>
<td>.008</td>
</tr>
</tbody>
</table>
CHAPTER VIII

Discussion of Results

Some Sample Descriptions

The kind of descriptions that the subjects gave and the kind of material that the judges had to work with can best be explained by presentation of a few samples. Excerpts from the protocols of two subjects will be presented. Subject 42 was an 18-year old sophomore in the College of Arts and Sciences, majoring in history. Subject 43 was a 19-year old sophomore in the College of Arts and Sciences, majoring in pre-law and political science. The type of response given and also some differences between subjects are illustrated by their descriptions of the young man in picture 6BM of the Murray series. In the protocols below, the words and phrases used to describe the figure are given first; then the antonyms of those words and phrases are given after them in parentheses. The words underlined indicate what the subjects thought to be the chief characteristics of the figure.

Subject 42 described the young man as

1. conscientious (unambitious)
2. loving (condescending)
3. rebellious (complacent)
4. intelligent (stupid)
Subject 43 described the same figure in the following way:

1. sadist, always griping (enjoyable)
2. serious minded (not serious)
3. disagreeable (agreeable)

The subjects gave all the descriptions of pictured figures in this form, although the particular content differed, in varying degrees, from one picture and one subject to another.

On the Repertory Test subjects were given three names at a time and instructed to sort the people and tell in what way two are alike and yet opposite from the third person. In sort number 7, for instance, they were instructed to use the names corresponding to the titles "your mother", "an older woman who is nearest your ideal", and "an older woman that you do not like". Subject 42 paired the first two people (i.e., first two titles above) and described them as "tactful", while he said the third person (third title, above) "lacks tact". Subject 43 grouped the people in the same way, but described the first two as "faithful" and the third person as "sly". In sort number 2, the names used corresponded with the titles "your mother", "your present or most recent girl friend", and "a male acquaintance about your own age that you do not like". Subject 42 grouped the first two people together and described them as "charming, practical,"
and intelligent but not intellectual", while he de-
scribed the third person as "unorthodox in manners, impractical, and very intellectual". Subject 43 grouped
the three people in the same way. He described the first
two as "comforting, congenial, and pleasant" and he
said the third person "makes others feel uneasy, is a griper and dissatisfied, and is moody".

When the protocols were prepared for the judges, the descriptions obtained from the picture test were
used in the same form, or almost the same form, as that in which they were obtained. Only a few changes, such
as spelling, grammar, suffixes, and word form (adjective or verb) were made in some words or phrases; these are
described in Chapter VI in the section on Preparation of the Materials for Matching. The material obtained from
the Repertory Test was reorganized, so that all adjectives and phrases describing any one person were placed to-
gether. In the example above, for instance, the title "your mother" is used in both sorts 2 and 7. The de-
scriptions of the subject's mother obtained from both sorts were grouped together, so that in the material by
Subject 42 given to the judges the description read

"charming (unorthodox in manners)
practical: (impractical)
intelligent but not intellectual (very intel-
tlectual)
tactful (lacks tact)"

(In the reorganized descriptions the words used to describe the person sorted as "opposite" are given in parentheses.)

The protocols from the two tests for each subject were broken up in various ways for matching; this has been explained previously. Part of the materials from one set used in testing hypothesis 4b can serve as an example and can illustrate some of the cues that the judges had to work with. Hypothesis 4b was selected because it presented one of the more difficult matching tasks.

In the test of hypothesis 4b each set of materials to be matched contained two groups of descriptions by each of four different subjects, but all the people described were ones commonly said to belong to a single age-sex group. The judges' task was to put together the descriptions given by the same subjects. In the illustration below, all the descriptions are from a set containing only descriptions of young men. Subject 42's descriptions of three young men (himself, a liked person, and a disliked person) obtained from the Repertory Test are listed below in the form in which they were given to the judges.

"introverted (extraverted)
definite goal (no definite goal)
unassuming (overly officious)
acting much of time (one personality: gossip)
The same subject described the three young men in the picture test in the following way:

emotionally upset, nervous (calm)  
intelligent (stupid)  
audacious (mundane, sophisticated)  
confused (well adjusted)  
deceitful (honorable)  
charming (artless)  
unreceptive to friendship (stimulating, magnetic personality)  
no definite goal (definite goal)  
unorthodox in manners (charming)  
impractical (practical)  
very intellectual (intelligent but not intellectual)  
very intelligent and intellectual (competitive personality)  
rebellious (complacent)  
intelligent (stupid)  
loving (condescending)  

Judge B matched these two protocols correctly, but Judge M did not. The judges did not state what cues they used, but some of the available ones can be discussed. (This should not be taken to be an exhaustive analysis, nor should it be assumed that only the types of cues used in this particular matching can be used in matching other protocols.) In this part of the Repertory Test the subject uses the dimensions "definite goal (no definite goal)" and "accomplished (does little)"; in the picture
test the same terms are not used, but a dimension which may be related is "conscientious (unambitious)". Both protocols show some interest in intelligence. People are described as "charming" and "deceitful", or those terms are used as antonyms, in both protocols. However, certain dimensions that appear in one test do not appear in the other (at least in the part quoted here), and if seemingly related ideas are found in the protocol of another subject in the same set of materials, a judge may make an erroneous matching. Apparently that is what happened when judge M matched the following descriptions from the picture test protocol of subject 43 with the Repertory Test protocol of subject 42. Subject 43 described the young men in the picture test as follows:

"worry wart (never worried)  
hard to get along with (congenial)  
quiet (boisterous)  
sadist, always gripping (enjoyable)  
serious minded (not serious)  
disagreeable (agreeable)  
emotional (unemotional)  
a fool (level headed)  
gullible (disbeliever)"

In subject 42's Repertory Test protocol are the dimensions "reserved (very extraverted)" and "conceited (unassuming, reserved)", while in subject 43's picture test protocol is the description "quiet (boisterous)". In the Repertory Test are the descriptions "definite goal (no definite
goal)" and "accomplished (does little)"; while the picture test contains the dimension "serious minded (not serious)". Furthermore, Repertory Test figures are described in terms of being "impractical (practical)", and the other subject describes a pictured figure in the terms "a fool (level headed)". Judge M's error in matching these two subjects' protocols might be called a "good error". Judge B's decision about the correct matching may have been influenced in this particular case by several appearances of the same verbal symbol on the two protocols. (The words "charming", "deceitful", and "intelligent" appear on both protocols.) Identity of verbal symbols cannot be used as the only cue, since sometimes two or three subjects in a set use some of the same terms, but it can be used in conjunction with other cues. Use of identical terms here shows the likely relatedness of dimensions that appear on the two tests, whereas Judge M's incorrect matching seems to have been based on his own abstraction and his own relating of the subjects' dimensions.

The particular records quoted above were chosen because they show more apparent similarities between subjects than do some records. In the same set with these two subjects' protocols were protocols by two other subjects. Both judges rightly differentiated between those protocols
and the protocols of subjects 42 and 43. As an example, the picture descriptions of one of the other subjects are given below:

*not friendly (friendly)
thinks of self mostly (friendly)
aggressive (easy)

serious (carefree)
slow (normal or fast)
intelligent (not well educated)

dishonest (honest)
carefree (serious)
likes to be alone (friendly)

Although this subject uses the term "intelligent", his antonym is "not well educated" whereas the antonym of subject 42 is "dull"; that is, the dimension appears to be different for the two subjects. The subject above uses the common terms "dishonest (honest)", while subject 42 uses the somewhat less frequent "direct manner (devious, deceitful)". It is in such ways as these that the differentiations can be made.

**Interpretation of Tests of the Hypotheses**

**Hypothesis 1:**

There are both perceptible relationships between certain TAT-type drawings and the descriptions elicited by those drawings, and perceptible distinctions between descriptions elicited by different drawings.

All three methods of statistical analysis indicate that the null hypothesis may be rejected, and that there
do exist significant and perceptible relationships between the drawings used in this experiment and the descriptions elicited by those drawings, along with distinctions between descriptions elicited by different drawings. In terms of personal construct theory this means that college students, at least, do not use constructs independently of the situation in which they are given. They make discriminations (differential predictions) between various situations. This is as we would expect of non-hospitalized subjects. An important number of the subjects, at least, have no one small set of constructs used in regard to all people within a given age-sex group. That is, the generality of at least some of the constructs used about figures within a given age-sex group is such that they do not take in all the figures in the age-sex group.

The finding that there are demonstrable relationships between descriptions and the figures that elicited them serves as a caution for the use of picture tests in clinical practice. If, for instance, a patient describes a pictured figure of an older man in a certain way, it cannot be assumed without test that he will describe other pictured figures of older men (and presumably, also, real people, such as his father) in just the same way. He may or may not do so. At most, the descriptions of one or a few pictured figures tell what dimensions are in a patient's
construct repertory and so what dimensions he may possibly use in other situations.

However, the fact that the judges made errors in approximately 38% of their matches suggests that the descriptions may not be related solely to the particular situations in which they are elicited. The finding lends plausibility to the existence of personal construct systems which the subjects bring into new situations.

**Hypothesis 2:**

There are both perceptible relationships between the ways in which any one subject describes several pictured figures commonly said to belong to the same age-sex group and perceptible distinctions between the ways in which he describes figures belonging to different age-sex groups.

Statistically, this hypothesis is in questionable status. Analysis of the total number of accurate matches by means of Vernon's modified contingency coefficient and by the critical ratio technique suggest that the null hypothesis should be rejected, although at a relatively low level of confidence. Analysis of the distribution of accurate matches by means of chi-square indicates, however, that the distribution is not significantly different from what might occur by chance.

The description used in this experiment do not appear to show many practically important similarities and dif-
ferences in terms of age-sex groups of the figures described. Apparently, for the majority of subjects tested, generality of constructs does not occur to any great extent along age-sex group lines. The majority of subjects, for instance, do not appear to have many constructs about young men as a group which are seldom used about young women, older men, or older women. Or, if they do have such constructs, their differential use is not evident to the judges.

The failure to reject definitely the null hypothesis may be interpreted in either of two ways. A choice between the two interpretations is dependent upon the results from tests of the other hypotheses. On the one hand, the descriptions of the three figures within any single age-sex group may differ considerably from each other, so that the judges can perceive no unity in the descriptions about figures within any one age-sex group. The descriptions of figures within a single age-sex group may be quite different from each other if they are a function, to a high degree, of the drawings themselves. The tests of hypothesis 1 have shown that there are perceptible relationships between the figures within any one age-sex group and the descriptions elicited by them. The question becomes one of whether that relationship between pictures and descriptions is sufficiently great to prevent the ap-
pearance of detectable relationships between the descriptions of three figures within a given age-sex group. On the other hand, the descriptions of figures belonging to different age-sex groups may be quite similar. That is, the generality of constructs may be so wide that their use cuts across age-sex group lines, and the judges cannot distinguish between descriptions on the basis of age-sex group. Clarification of both of these possibilities is given by the results of hypotheses 3b and 3c, which are discussed below. Regardless of which possibility appears most likely, there is an indication, in the failure to reject the null hypothesis, that generality of personality constructs does not occur in terms of age-sex group.

Although the experimental hypothesis can be neither rejected nor accepted with confidence, whatever generality of constructs may exist in terms of age-sex group is likely to be of little practical use in work with individuals. If a high level of statistical significance cannot be shown with 32 subjects, the relationships tested are not likely to be evident in the examination of many of the subjects as individuals.

Hypothesis 3a:

There are perceptible relationships between the descriptions of various pictured figures given by any one subject when those figures are grouped into two sets, each composed of figures representing several age-sex groups, provided that the age-sex groups are
the same in the two sets, and there are perceptible distinctions between the descriptions of the same figures given by different subjects.

In terms of all three methods of statistical analysis the null hypothesis can be rejected with a high level of confidence. Although the sample was small, involving only eight sets with a total of 32 pairs of protocols to be matched, 97% of the total matchings by the two judges were correct. Judge B matched all 32 pairs accurately, and Judge M matched 30 of 32 pairs accurately.

In this test, the entire picture test protocol of each subject was divided into two parts, and it was the two parts that the judges matched. The results indicate that, at least in this sample of subjects, individuals do show some similarities in the way they describe various pictured people, and that different subjects describe the same people in somewhat different ways. Since both parts contained descriptions of figures in all four age-sex groups, no light is shed on the variable of age-sex group as a factor in generality of constructs. All that can be said here is that subjects do show some generality of constructs from one situation to another within the same test. Further analysis of areas of generality is given in terms of hypotheses 3b, 3c, 4a, and 4b.
Hypothesis 3b:

There are perceptible relationships between the descriptions of various pictured figures given by any one subject when those figures are all commonly said to belong to the same age-sex group, and there are perceptible distinctions between the descriptions of the same figures given by different subjects.

Tests of this hypothesis involve only one-fourth the amount of material used in testing hypothesis 3a, for in each set of materials to be matched, there were descriptions of figures in only one age-sex group. The three descriptions by each of four subjects were divided into two parts, and the judges' task was to put those parts back together again. Analysis of the matching results lead to rejection of the null hypothesis at better than the .0001 level by the contingency coefficient and critical ratio methods, and rejection at better than the .001 level by the chi-square method (for which only 3-place tables were available). This indicates that there are perceptible similarities between the ways that individual subjects describe different figures belonging to the same age-sex group, and that different subjects vary in their descriptions of the same figures. Although matching for this hypothesis was much better than chance matching, it was not so accurate as for hypothesis 3a; only 71% of the matchings were correct, as opposed to 97% for hypothesis 3a.
The difference is probably based on the fact that for hypothesis 3b the judges had, at any one time, only one-fourth the amount of material available for hypothesis 3a. Because of the smaller sample of material from each subject, there is less chance for generality of constructs to show up.

Rejection of null hypothesis 3b suggests that the failure to reject null hypothesis 2 was not based on complete, or nearly complete, lack of generality of constructs within given age-sex groups. The descriptions of the three figures within any age-sex group cannot be said to differ entirely from each other and to be based exclusively on the various aspects of the drawings that elicited the descriptions. The alternate interpretation of the failure to reject hypothesis 2 must be given consideration, and this will be done in the discussion of the results of the test of hypothesis 3c.

**Hypothesis 3c:**

There are perceptible relationships between the descriptions of various pictured figures given by any one subject when those figures are of different age-sex groups, and there are perceptible distinctions between the descriptions of the same figures given by different subjects.

Analysis by both the contingency coefficient and critical ratio methods indicates that the null hypothesis may be rejected at less than the .0001 level of confidence,
while analysis in terms of chi-square indicates rejection at less than the .001 level, which is the best that available tables allowed. There is demonstration that, in this group of subjects, individual subjects described figures belonging to different age-sex groups in similar ways and that, also, the descriptions of the same figures by different subjects varied. The high degree of accuracy with which the materials could be matched is shown by the fact that in twenty-four sets of materials, with a total of 96 pairs of protocols, Judge B matched 94 of the 96 pairs correctly, and Judge M matched 87 of the pairs correctly. (Probably the material was matched more accurately than the material for hypothesis 3b, because each test of hypothesis 3b involved twice as many descriptions as did each test of hypothesis 3c.) The accuracy of matching suggests that the great majority of subjects show perceptible generality of constructs that is not limited within the areas of single age-sex groups. Generality of constructs appears to be either quite broad or to occur in terms of other than age-sex groups, for the judges were able to match descriptions of young men with descriptions of young women, older men, and older women. This is not to say that there are no subjects whose constructs are defined by the age-sex areas, nor that the majority of subjects have no constructs defined by age-sex limits. It
does mean that an important number of many subjects' constructs are not limited in those terms.

Rejection of null hypothesis 3c and failure to reject null hypothesis 2 are quite consistent. Findings in reference to hypothesis 3c suggest that matching for hypothesis 2 was not more accurate because descriptions about figures in different age-sex groups resemble each other so much that the judges could not distinguish one group from another. The judges' spontaneous comments during the training session also indicate difficulty in distinguishing between the groups. The results of these two hypotheses suggest that college men have one or more sets of constructs whose generality extends over people differing in either or both age and sex. If these findings are confirmed with clinic patients we might suggest that less stress need be put on the variables of age and sex, than is sometimes done, in regard to patients' relations with other people.

Hypothesis 4a:

There are perceptible relationships between the descriptions of pictured figures and real people given by any one subject when the people in both tests belong to several age-sex groups, provided that the age-sex groups used in the two tests are the same, and there are perceptible distinctions between the descriptions of the same figures given by different subjects.

All three methods of analysis indicate that the null hypothesis may be rejected, although the contingency
coefficient and critical ratio methods do so with a higher level of confidence than the chi-square method. The matching results indicate that individual subjects show some similarities in the way they describe pictured people and real people in two different test situations, and, also, that the descriptions of different subjects vary. They suggest that the generality of constructs used about pictured figures is broad enough to embrace real people, also. In this test, as in the test of hypothesis 3a, all four age-sex groups were used, and so nothing can be said about generality in terms of the age-sex areas.

Hypothesis 4b:

There are perceptible relationships between the descriptions of pictured figures and real people given by any one subject when the people in both tests are all commonly said to belong to the same age-sex group, and there are perceptible distinctions between the descriptions of the same figures given by different subjects.

Here, also, the null hypothesis is rejected by all three methods of statistical analysis. Rejection indicates that individual subjects do show similarities in the way that they describe pictured and real people in two different test situations when both sets of people belong to a single age-sex group, and rejection also
indicates that the descriptions of various individuals differ somewhat from each other. The results of this test give further evidence, in addition to that given by hypothesis 4a, for generality of constructs that is broad enough to include both pictured people and real people known to the subjects. It cannot be said whether the generality is broad enough to extend across age-sex lines as well as across type of test, for no hypothesis analogous to 3c was tested. The accuracy of matching was lower for this hypothesis than it was for hypothesis 4a, probably because each set of descriptions to be matched contained only one-fourth the amount of material that the sets contained for hypothesis 4a. Thus the opportunity for generality of constructs to appear was reduced in comparison with the opportunity in hypothesis 4a. For hypothesis 4b, 32 sets including 128 pairs of descriptions were matched, and 47% of the total matchings by the two judges were correct. Judge B matched 71 pairs correctly, and Judge M matched 50 correctly. The fact that even this degree of accuracy was shown for the small sample of material increases the probability that constructs elicited by picture tests may be used in other situations, even non-test situations. Picture tests as a means of eliciting subjects' constructs in clinical settings will be useful, of course, only if that kind of broad generality can be demonstrated.
Area of Permissible Application of Results

The three major variables of subjects, test materials, and judges bring up the question as to the area of permissible application, or generalization, of the experimental results. Clearly we cannot expect future judges necessarily to do the same kind of matching that these judges did. The sample of two judges is much too small for such a conclusion. It is too small for us to expect even other trained judges to do the same kind of matching, and it is trained judges that we are interested in. The function of the judges in this experiment is to serve as a kind of measuring instrument; the judges themselves are not being studied. As instruments, their matchings varied so much from chance matchings that they gave us information about the characteristics of the materials. But they did not give us much information about the characteristics of judges. For that kind of information, a much larger number of judges would have to match the same materials, and the central tendency and variability of the judges would have to be determined.

Neither can we conclude from these experimental results that the use of other pictures or other forms of the Repertory Test would lead to the same results. The sample of test materials is much too small for that.
The major interest in this experiment is in the constructs elicited by the pictures rather than in those elicited by the Repertory Test, and consideration of the wide variation possible in pictures suggests that a very large sample, indeed, would be needed before any broad statements could be made about constructs elicited by them. The sample would have to be much larger than the sample of judges mentioned above, for the variability of judges is reduced by training, but the variability of pictures is almost limitless.

In the area of subjects, we cannot assume that all subjects will give descriptions suggesting the same types of generality as did the subjects used here. Because of the highly significant test results it is likely that other college men would show similar generality. It is also possible that other, non-college subjects would give responses leading to similar experimental results, if they described the same figures, and if their descriptions were matched by the same judges; however, this must be determined experimentally. Although the experimenter's interest in the problem of generality began in relation to clinic patients, the present results cannot be applied to them unquestioningly. Further studies are needed.
Observations about the Judges

The accuracy of matching is affected by the characteristics of the judges as well as by the characteristics of the subjects and the materials to be matched. During the training session both judges entered actively into the tasks and put a good deal of effort into their attempts to discover clues for matching. However, there were two major differences between the two judges. One difference between them was in previous experience with personal construct theory. Judge B had had some familiarity with it whereas the approach was completely new to Judge M. However, an introduction to the theory was given Judge M during the training sessions. As nearly as could be determined by observation, familiarity with the theory was most important in use of dimensions rather than merely descriptive words as cues in matching; Judge B made more use of dimensions during the first training session than did Judge M, but Judge M came to use dimensions more during the second training session than he had during the first session. Another difference between the judges occurred in their approach to the matching. Their behavior during the training sessions, as well as casual observations on previous occasions, suggest that Judge M's efforts probably varied less from one part of the task to another, but that on the whole he took the task more casually than
did Judge B. Judge B's usual approach appeared to be more analytical and somewhat more careful than Judge M's, but if she felt little sense of success and if the task appeared to be impossible after reasonable effort, she made her final selections quickly and almost haphazardly.

On six of the seven hypotheses and sub-hypotheses Judge B achieved a greater accuracy than did Judge M. Sometimes the difference was only a small fraction of the total number of matches made (e.g., a difference of 3 out of 128 pairs for hypothesis 2), but occasionally the difference was relatively large (e.g., 21 out of 128 pairs matched for hypothesis 4b). Perhaps most of the difference that existed was due to a combination of Judge B's greater familiarity with the theory plus her usually more analytical work habits. On the other hand in the matching for hypothesis 4a Judge M was more accurate, matching 22 out of 32 pairs, while Judge B matched only 14 out of 32 pairs. This hypothesis required the use of the complete protocols from both tests, and this was more material than any of the other hypotheses involved. During the training sessions Judge B appeared to feel more frustration than did Judge M over the difficulty of keeping in mind so much information at one time, and when she became frustrated, the care with which she matched decreased. It is possible that this sort of thing happened during the experimental matching.
Despite the differences between the subjects, they showed significantly better than chance agreement with each other in the matching for most of the hypotheses. The percentage of identical matches by the two judges varied from the low 40's for the most difficult tasks (hypotheses 2 and 4) to the high 80's and low 90's for the easiest tasks (hypotheses 3a and 3c). On hypotheses of intermediate difficulty (1 and 3b) their agreement was in the 60's. Chi-square tests of the difference between their patterns of accuracy of matching (in terms of number of pairs correct per set of materials) were non-significant for most of the hypotheses, but they did show significant differences in pattern on hypotheses 3b and 4b. These were the hypotheses involving the smallest amounts of material per subject. When they were given only a small amount of material by each subject, Judge B apparently could make keener judgments about the material and therefore could obtain a greater number of sets in which all or most of the pairs of descriptions were matched correctly. In most cases, because the two judges did not show significant differences in either total number of pairs matched correctly or in pattern of accuracy of matching, the mean of their results was used. In the case of hypotheses 3b and 4b, however, because they did show significant differences in pattern, the results of the two
judges were analyzed separately as well as in terms of their mean results. In all the analyses the results of both judges were greatly different from chance results, although the results of Judge B showed a higher level of significance.

Various measures of similarity of judges have been reported in the literature, and the obtained degrees of similarity between the judges vary widely. Variability with difference in measure used is shown in the report of Allport, Walker, and Lathers (5). Their two judges both averaged four out of a possible eight correct matchings per subject, but they varied considerably as to which subjects they identified most correctly, and the correlation between the identification indices was only -.28. The authors comment that the reasons used for matchings varied, and that the matchings showed the ability, background, and traits of the two judges as well as the characteristics of the subjects. Although the same type of index was not calculated in the present experiment, the percentage of agreement was calculated, and this is somewhat comparable. In terms of the findings of Allport, et al, the percent of identical matches obtained here appears relatively high.

Some of the previous studies report use of both "experts" and "laymen" as judges, usually with considerably
better matching by the "experts" (16, 36). Some experi-
menters, such as Secord (107) have used only laymen, and
some, such as Palmer (83), have used only so-called ex-
perts. The present study is the latter variety. If the
major interest is in the material or the subjects, rather
than in the judges themselves, the judges used should be
the best ones available. This is what Allport (5) concluded
should be done.

Even among so-called experts, experimenters have found
considerable variability. Murray, for instance, reports
that when four judges who were intimately acquainted with
5 subjects attempted to identify test responses, one judge
was successful three times, one judge two times, and two
judges only once each (31:279). The present results are
not directly comparable with Murray's study, for the pre-
sent ones appear to show much greater similarity between
the judges, since in most cases the differences between them
are non-significant, while the difference between one and
three correct matches would probably be significant. (Of
course, to be sure, about the similarity of the judges more
matching tasks are needed in Murray's experiment). Per-
haps the greater similarity is due to the preliminary train-
ing and definition to the judges of the bases of matching.

In addition to the factor of the statistical measure
used and the ability and training of the judges, another factor affecting the comparability of judges from experiment to experiment is the homogeneity or heterogeneity of materials to be matched. If the materials are heterogeneous and matching is easy, then the judges in a single experiment may be expected to show considerable similarity to each other. But if the materials are relatively homogeneous and the matching task is difficult, then the judges are likely to differ more from each other because they may sometimes be "grasping at straws" and may grasp at different straws. Heterogeneity is a factor that has been treated in various ways by various experimenters, and the different treatments reduce the comparability of judges' reliability just as they reduce the comparability of the matching results themselves. Strictly speaking, the judges in the present experiment should be compared only with judges in experiments in which the materials were randomized.

Taken altogether, the factors of randomization of the materials, type of materials to be matched, amount of training for the judges, and type of measure of comparability of the judges mean that there is no experiment directly comparable to the present one, and, therefore, no set of judges with whom the present judges may be compared.
Observations about the Subjects

Most subjects required about 45 minutes to complete the picture descriptions and associated tasks, and about 45 to 50 minutes to complete the Repertory Test. The time for individual subjects varied widely, however. For the Repertory Test some subjects needed only 30 minutes and one subject needed 70 minutes, while one subject apparently could not do the test at all. For the picture descriptions and associated tasks some subjects required only 35 minutes, but one subject required 95 minutes.

Behind the variability in time required were wide differences in other characteristics of the subjects. An important number of subjects (the experimenter estimates about one-quarter of the entire group tested) stated that they had never thought about the personality characteristics of people, particularly people they knew fairly well, and their first descriptions were in terms of such variables as sex, age, physical characteristics, occupation, recreational interests, or specific actions. Some subjects had great difficulty getting off this level of description, particularly when describing real people on the Repertory Test. About one-sixth of the students tested seemed to base their descriptions of pictured people almost entirely on the faces of the people; and so they had
great difficulty whenever faces were not shown clearly, as, for instance, on Murray card 12M. Limited vocabulary appeared to be a problem for some subjects, so that they tended to use long phrases or whole sentences to explain their ideas. In the language of personal construct theory, these subjects did not appear to have sufficient verbal symbols for their constructs. About half the subjects considered the two tasks quite difficult for one or another of these reasons.

It was necessary to discard and not use as subjects a much larger number of students tested than the experimenter had anticipated. She had hoped that careful selection of the pictures would mean that almost all the students would agree on the age-sex group identification of all the pictured figures and that consequently it would be necessary to discard only a small percentage of the students tested. The total proportion of misperceptions of figures was small, but the fact that most students who misperceived any figures misperceived only one figure meant that the number of students who were dropped was fairly large (one-third of those tested). The number of students who showed misperceptions brings up the question of the significance of the misperceptions. If one or two misperceptions are a sign of psychological difficulties, as they are often taken to be, then there is a suggestion of fairly widespread difficulties
among college men. On the other hand, perhaps the misperceptions are not so serious a sign of maladjustment as they are sometimes said to be, and their use in clinics should be modified. Eron (24) showed that in regard to picture 3 of the Murray set, misperception of the sex of the figure is probably not important, but perhaps less importance should be attached, also, to an isolated misperception of another figure. Further investigation of the identification of figures is needed.

Some subjects, it was noted, used a single pattern for grouping people on all the sorts of the Repertory Test; that is, they always grouped the liked person and the "principal" person (girl friend, parents, or subject himself) together as being alike and placed the disliked person alone as being different from the other two. Other subjects varied their pattern of grouping. An hypothesis might be set up that descriptions of the former subjects will be more global and evaluative than the descriptions of the latter subjects. The descriptions of the latter subjects, on the other hand, might be expected to be more analytical and descriptive than the descriptions of the former subjects. Apparently the latter subjects were making a real effort to describe the people, while the former subjects were merely saying in various ways that they liked or disliked certain people. It seems possible that
in many cases the terms applied to the people by the former subjects were merely labels applied to anyone that the subjects liked or disliked; in the case of the disliked people, the supposedly descriptive terms appeared to be a kind of disguised name-calling. Use of this or a similar form of the Repertory Test, with attention to the pattern of grouping as much as to the constructs elicited, might be one aid in evaluating subjects' potentialities for jobs in which an important part involves considerable discrimination in their treatment of various people. For the subjects who use the more analytical descriptions and who use variety in pattern of grouping appear likely to show more differentiations in their actions toward different people and toward a single person at various times.

In another respect, also, subjects showed differences from each other, while at the same time remaining fairly consistent with their own pattern. It is estimated that about two-thirds of the subjects began their sorting at the likeness end of the constructs, describing in what respect two of the three people were similar, and then afterwards stating the way that the third person was different. About one-third, however, stated first how one subject differed from the other two, and then stated afterward how the other two were the same. This latter proce-
dure was actually contrary to the request made in the test directions to "tell in what way two of the people are the same and yet different from the third person", and therefore it may be of more diagnostic significance than the former procedure. Although this form of the test might suggest that that procedure is related to the division into liked and disliked groups, the fact that the same procedure appeared on earlier and different forms of the test suggests that that is not the primary relationship. Perhaps those subjects make few differentiations between people, seeing nearly everyone as being described by the same constructs. Then if all people look much the same, they may not have many terms readily available for describing the sameness. They may have terms readily available for describing only people who do not fit into their idea of people as a whole; then they may have to ask themselves, "What is the opposite of _____?" in order to describe the people whom they see as the same. The judges in this experiment, like the judges in the experiment of Allport, Walker, and Lathers (5), commented that some subjects' protocols were much easier to match than other subjects' protocols. The analysis of matching for the first hypothesis lends support to the subjective impression of the judges. Actually, in a few cases, the accuracy of matching some subjects' protocols was twice the
accuracy of matching other's protocols. A statistically significant difference between subjects falling in the upper and lower quartiles was shown. In the case of hypothesis 1, the difference appears to indicate that some subjects were heavily influenced by the picture content in giving their descriptions of various figures. Other subjects appeared to show relatively little difference between their descriptions of various figures, and presumably they brought into each situation constructs that were related relatively little to the particular situation. It seems likely that the latter subjects show greater generality of constructs than the former subjects.

Tests of hypotheses 3 and 4 involved protocols of four subjects in each set of materials, and for these hypotheses, also, the judges reported differences from one set to another in ease of matching. Confirmation of their subjective experience comes in the fact that the judges certainly showed variation in accuracy of matching from one set to another. Such differences might be based on differences in either the generality or personal quality of constructs used. If one individual's descriptions do not differ much from others' descriptions, then his cannot be distinguished from the other person's descriptions. This is the problem of homogeneity-heterogeneity of subjects. Also, even though an individual's protocols are
distinguishable from the protocols of other people, if they do not show some relationships between descriptions of various figures, the two parts of his records cannot be matched. Both a personal quality and generality of constructs are necessary for successful matching much above a chance level.

It was thought originally that there might be negative correlations between the accuracy with which subjects' protocols could be matched on tests of hypothesis 1 and on tests of hypotheses 2, 3, and 4. It seemed reasonable to expect that if subjects' descriptions of figures were perceptibly related to the particular pictures that elicited the descriptions, there might be relatively little relationship between the three descriptions of figures in any given age-sex group, and probably, also, relatively little relationship between the descriptions of figures belonging to different age-sex groups. On the other hand, it also seemed reasonable to expect that if there were no perceptible relationships between the descriptions of figures and the pictures that elicited them, then there might be considerable similarity between the descriptions of three figures in any given age-sex group, and, perhaps, also, some relationship between the descriptions of figures belonging to different age-sex groups. However,
correlations of the results of hypothesis 1 with the results of hypotheses 2, 3, and 4 showed no statistically significant relationships, either negative or positive. The lack of any significant correlations, along with positive results for hypothesis 1, as well as for hypotheses 3 and 4, suggest that the descriptions (each of which consists of three or four adjectives or phrases) contain some aspects that are closely related to the pictures that elicited them, and, in addition, some aspects that are personal in nature and that appear both in descriptions of various figures within a given age-sex group and in descriptions of figures belonging to different age-sex groups.

Originally it was also thought that the accuracy of matching individual subjects' protocols might be positively correlated in tests of hypotheses 3 and 4. Accurate matching for each of those hypotheses requires both a personal quality in the descriptions by each subject and some relationship between descriptions of different figures by any given subject. In addition, hypothesis 4 requires relationships between descriptions given to different kinds of tests, and the chances that these will be found are increased if there are perceptible relationships between descriptions of figures appearing in
a single test. A correlation of .436, significant at the .008 level of confidence, was found between the "matchability" of subjects' protocols in tests of hypotheses 3 and 4. Apparently the subjects whose constructs were general enough to be applied to various pictured figures also tended to show generality of constructs over an area broad enough to include both real and pictured people; and those whose descriptions of pictured people did not show much generality of constructs (i.e., those whose constructs were relatively specific to particular pictured figures) also showed little use of the same or similar constructs in regard to both pictured and real people.

A correlation of -.379, significant at the .017 level of confidence, was found between the "matchability" of subjects' protocols in tests of hypotheses 2 and 4. This correlation was not unexpected. Apparently, when the judges found it difficult to distinguish between a subject's constructs about figures in different age-sex groups, they found it relatively easy to distinguish his constructs from the constructs of other people and to match his constructs about pictured and real people if the age-sex group membership of the figures was not important. Consistent with the negative correlation
between the results of hypotheses 2 and 4 was a low negative correlation of borderline significance between the results of hypotheses 2 and 3. When the judges found it difficult to distinguish between a subject's constructs about figures in different age-sex groups, they found it relatively easy to distinguish his constructs from the constructs of other people and to match his constructs about various pictured people if the age-sex group membership of the figures was not important.

Before the testing was begun, the possibility was considered that many subjects might try to give socially acceptable answers, saying only good things about the people, and suppressing their first (silent, unacceptable) responses. It is impossible to measure accurately or to eliminate this tendency. It seems likely that if the tendency appeared at all, it would be stronger on the Repertory Test than on the picture test, because in the former test the subjects are asked to describe people they know, whereas in the latter test they describe fictitious characters and can rationalize their descriptions by saying that characteristics of the pictures provoke them. On the Repertory Test, an attempt is made to minimize the tendency (1) by the use of a modified forced choice technique, (2) by statements to the sub-
jects that their names will not be placed on the test records, that no report will go to their class instructor or to anyone else, and that the analysis of results will be in terms of groups rather than in terms of individuals, and (2) by allowing the subjects to keep the names of the people hidden if they wish. During testing a few of the subjects did show reluctance to make any unfavorable statements about the people on their name lists, but most showed no such hesitation and actually made a number of unfavorable statements. If suppression of the most available, and socially "unacceptable", responses did occur on either test, the amount of demonstrated generality of constructs probably would be reduced. Then actual generality of constructs used in reference to pictured and real people probably would be greater than was shown, rather than less than was shown. If suppression of the most available, and socially "unacceptable", responses did occur on both tests about equally, the demonstrated generality of constructs would be affected in an unknown direction and degree; if the subjects merely used the contrasting ends of the construct dimensions (the most likely response, according to personal construct theory) so as to give "acceptable" responses, then there would probably be little influence on the demonstrated generality of constructs.
Observations about the Picture Test

Each of the twelve pictured figures was chosen for its supposed ability to elicit a wide variety of descriptions and for its supposed identification by nearly all subjects in terms of a single age-sex group. The comparative ambiguity of the three pictures in each age-sex group was studied by means of chi-square. As has been stated previously, if the figures were all perfectly ambiguous, and if all descriptions were matched by chance alone, then one-third of the correct matches for each age-sex group should involve each picture in the group. The obtained distribution of correct matches for each age-sex group was compared with the expected distribution. No significant difference was found between the two distributions for the older male, younger male, and younger female groups. All three pictures in each of these groups appear to be about equally ambiguous or unambiguous. Within each of these three age-sex groups, therefore, all three pictures appear to be about equally useful for eliciting personal constructs. The distribution of correct matches within the older female group deviated significantly from that expected by chance alone. Murray picture 5 was matched more accurately than either of the other two pictures. This it appears that that picture is less ambiguous or provokes a smaller variety of responses or more stereotyped responses.
than the other pictures in the same group. If the constructs elicited by picture 5 are closely related to the picture itself, then that picture is likely to prove less useful than the other two pictures in eliciting personal constructs which may be used over a broad area.

If the pictures in the four age-sex groups are equally ambiguous, and if the descriptions are matched by chance alone, then one-fourth of the total number of correct matches should be made in reference to each group. When the pictures of the four age-sex groups are compared as groups, it is seen that pictures of the young females were most often correctly matched with their descriptions elicited by them and that pictures of the young males were least often correctly matched with their descriptions. The distribution of obtained correct matches for the four groups differs significantly from that to be expected by chance. Apparently the descriptions of the young females show both greater differences between themselves and greater relationships to the pictures than is true of the descriptions of the young males and also of the descriptions of the older females. In terms of eliciting personal constructs that are likely to be used in a wide variety of situations, the young male pictures and two of the three pictures of older females are probably more adequate than
the pictures of young females and the pictures of older males.

In this study it was desired, for clinical reasons, to utilize published pictures that are widely known and used. It was recognized that the published pictures might not all be equally useful for eliciting personal constructs. But suppose that instead of these pictures, some pictures had been used that were especially drawn to be ambiguous. It is expected that the latter pictures would elicit a higher proportion of personal constructs having a relatively wide generality than would the former pictures. In using the present pictures, the experimenter was probably reducing the chances of demonstrating generality of constructs. Actually it is not surprising that some of the pictures elicited descriptions that were fairly easily identified, for the pictures were originally chosen by Murray and Symonds for the possibility that subjects might relate them to frequent problem situations, and the problem situations were chosen largely in terms of psycho-analytic theory. Here is a case in which different theories call for different test materials and procedures.

A survey was made of the figure identifications by both the subjects used and the students who were rejected as subjects for the matching study. Table XIII contains
summarizes of the figure identifications in terms of ages of the figures. In regard to the subjects used, it can be seen that although there is some overlap in ages assigned to the older and younger groups, there is no overlap of median ages. The overlap in range is possible because not only the assigned ages of the figures but also

<table>
<thead>
<tr>
<th>Figures</th>
<th>Subjects Used</th>
<th>Subjects Not Used</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range  Median</td>
<td>Range  Median</td>
</tr>
<tr>
<td><strong>Older Males:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12M-R</td>
<td>45-70  60</td>
<td>25-75  55</td>
</tr>
<tr>
<td>7BM-L</td>
<td>50-73  60</td>
<td>50-75  57</td>
</tr>
<tr>
<td>A4-R</td>
<td>28-55  39</td>
<td>30-65  40</td>
</tr>
<tr>
<td><strong>Older Females:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>35-69  50</td>
<td>40-65  50</td>
</tr>
<tr>
<td>A7-L</td>
<td>29-50  40</td>
<td>18-55  45</td>
</tr>
<tr>
<td>6BM-L</td>
<td>45-78  60</td>
<td>52-72  60</td>
</tr>
<tr>
<td><strong>Younger Males:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12M-L</td>
<td>15-26  18</td>
<td>16-50  22</td>
</tr>
<tr>
<td>6BM-R</td>
<td>23-40  30</td>
<td>25-34  30</td>
</tr>
<tr>
<td>13MF-R</td>
<td>19-38  25</td>
<td>19-55  28</td>
</tr>
<tr>
<td><strong>Younger Females:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-R</td>
<td>20-38  26</td>
<td>25-30  25</td>
</tr>
<tr>
<td>A10-M</td>
<td>14-23  19</td>
<td>15-35  20</td>
</tr>
<tr>
<td>13MF-L</td>
<td>17-32  24</td>
<td>18-26  25</td>
</tr>
</tbody>
</table>
the relationship between the ages of the two (usually) figures within any one picture were considered when the decisions were made about whether to retain or discard subjects whose identifications made them fall on the borderline of acceptability. Such borderline identifications occurred when either the so-called older or younger figure was stated to be about 35 years old. The criterion was set up that a subject would be retained if he maintained the common age relationship between the two figures even though he described either a so-called young figure or a so-called older figure as being about 35. If he did not maintain the age relationship and, e.g., described both "young" and "older" figures as being about 35, he was not used as a subject. For subjects used in the matching study, the median ages of the three older male figures were 60 years in two cases and 39 years in the other case, while the median ages of the three older females were 40, 50, and 60. The range in age assignments of older males was from 28 to 73, and the range for older females was from 29 to 78. This wide range indicates marked interpersonal differences in perception. For any one figure in the two older groups the age span covered varied from a 23 year span to a 34 year span. For the so-called younger males the median
ages of the three pictured figures were 18, 25, and 30, and for the younger females the median ages were 19, 24, and 26. The range in age assignments of younger males was from 15 to 40, and the range for younger females was from 14 to 38. For any one figure in these two younger groups the age span covered varied from a 9 year span to a 19 year span. Again marked interpersonal differences in perception are shown. There was a tendency for subjects who identified the younger figures at the lower end of the age range to identify the older figures at the lower end of their age range also, and vice versa.

When the identifications made by subjects who were rejected for the matching study are examined, it is found that the age range was wider for some pictures and narrower for others than in the case of subjects who were used in the matching study. The broader age ranges are often directly related to the rejection of the subjects on the basis of misperception of age-sex group. The reason why the age range was narrower for some pictured figures is not immediately evident; perhaps the variations are only chance fluctuations. The narrower age range was most noticeable in regard to the young man in picture 6BM (range from 25 to 34 years), the young woman in picture 4 (range from 25 to 30 years), and the young woman in
picture 13 MF (range from 18 to 26 years). The mean ages assigned by rejected subjects resembled the means assigned by the group of subjects used. For the various figures, the means of the two groups of subjects differed by from zero to five years for the figures in the older age-sex groups and by from zero to four years for the figures in the younger age-sex groups.

Seven of the misperceptions by rejected subjects were in terms of the age of the figure, ten were in terms of the sex of the figure, and two were in terms of both the age and the sex of the figure. All changes of sex were perceptions of males as females. Sometimes young figures were perceived as older, and sometimes older figures were perceived as younger. Twelve of the nineteen misperceptions occurred in regard to the young men on either picture A4 or picture 12M, whereas from one to three misperceptions occurred in regard to the older men on pictures A4 and 12M, the older woman on picture A7, and the young man on picture 13MF. The number and type of misperceptions are detailed in Table XIV.

When a tabulation was made of the social roles assigned to the various figures by the subjects who were used, it was found that the number of different social roles assigned to any one figure varied from five (for the
older woman in picture 6BM and the younger man in the same picture) to seventeen (for the older man in picture 12M). In the case of pictured figures it seems possible that the social roles assigned to the figures are a resultant of the constructs about the figures rather than that the constructs about the figures are a resultant of the social roles assigned to them. It may be that as a person looks at a figure, he construes it in some way that is consistent with his construction system as it has evolved in the past; then either spontaneously or upon request, he searches for some social role that he sees as fitting the constructs that he has used in regard to the figure. This hypothesis is put forward partly on the basis of the fact that many subjects hesitated when asked to assign social roles to some figures, even though they had previously described the "personality" of those figures. Another basis for the hypothesis is that within limits (as shown by the results of hypothesis 1) individual subjects appear to have one or more broad construction systems which are applied to figures in various age-sex groups (as shown by tests of hypotheses 2 and 3b). In view of this type of demonstrated unity of construct systems it seems unlikely that subjects would have sets of constructs that are even more narrowly used and are applied to specific social roles, except perhaps in a
relatively few cases. In terms of personal construct theory it might be said that the subject makes his predictions about the figures in terms of some characteristics.

TABLE XIV

Number and Type of Misperceptions of Pictured Figures

<table>
<thead>
<tr>
<th>Figures</th>
<th>Types of Misperception</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age</td>
</tr>
<tr>
<td>Older Males</td>
<td></td>
</tr>
<tr>
<td>12M-R, A4-R</td>
<td>1</td>
</tr>
<tr>
<td>Older Females</td>
<td></td>
</tr>
<tr>
<td>A7-L</td>
<td>2</td>
</tr>
<tr>
<td>Younger Males</td>
<td></td>
</tr>
<tr>
<td>12M-L, A4-L</td>
<td>0</td>
</tr>
<tr>
<td>13MF-R</td>
<td>3</td>
</tr>
<tr>
<td>Younger Females</td>
<td>None</td>
</tr>
</tbody>
</table>

of the drawings and his previous construct system together, and that the social roles assigned to the figures are really summaries of his predictions about figures, as if he asked himself, "who is such-and-such a kind of person?" In regard to real people we know that people are sometimes identified, for instance, in terms of guessed occupation. We have heard people say, e.g., "She looks like a schoolteacher," or "He looks like a farmer." Where strangers are identified in this way, the process seems
to be similar to the process in dealing with pictured figures. On the other hand, where people are introduced first in terms of social role, it may be expected that constructs summarized by the social role will come to the fore. This latter process may be said to be deductive, while the former one may be called inductive.

Earl}'' in the period of designing the present research the possibility was considered that the experimenter might name the social roles of the figures so that all subjects would use constructs about the same social roles. The idea was abandoned in favor of giving subjects a greater amount of freedom in an effort to elicit their more available constructs. It seems likely that if social roles were assigned, the constructs elicited often would be compromises between ideas about the social roles named and ideas (perhaps different) about figures drawn like the pictured figures.

The testing procedure was set up originally with the hope that some observations might be made on possible relationships between physical similarities and differences between drawings and the "matchability" of the descriptions elicited by the drawings. The eight preliminary subjects on whom this part of the procedure was tried out appeared to understand the questions about the drawings, and they gave
fairly clear-cut answers. Among the experimental subjects, however, many had considerable difficulty in understanding the questions and separating the physical characteristics of the drawings from their interpretations of the personalities of the figures. Because of their difficulty and because this part of the procedure was not an integral part of the tests of the hypotheses, several forms of questioning were tried out during the course of testing. Since no one standard procedure was used and since some subjects never did differentiate between the drawings and their interpretations of the drawings, the material about characteristics of the drawings was not analyzed. What this part of the procedure did not demonstrate was the wide between-subject variation in awareness of use of interpretations. Subjects having this awareness may be said to see themselves as construers of situations, and subjects without this awareness may be said to fail to see themselves as construers.

Observations about the Repertory Test

Let us now turn to a discussion of the Repertory Test materials and procedure. Because of the fact that relationships between descriptions were being determined for some of the hypotheses in terms of age-sex groups, tabulations were made of the ages of people whose names were substituted for the various titles. It was desired to see
whether the subjects differentiated much between the groups (as set up by the experimenter) in regard to age, and whether the ages of the people in the Repertory Test groups were comparable to the ages of the people in the picture test groups. The ranges and ages of people substituted for each title are summarized in Table XV. For both the subjects used and the subjects rejected, the

TABLE XV

<table>
<thead>
<tr>
<th>Titles</th>
<th>Subjects Used</th>
<th></th>
<th>Subjects Not Used</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Range</td>
<td>Median</td>
<td>Range</td>
<td>Median</td>
</tr>
<tr>
<td>Older Males</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>26-70</td>
<td>45</td>
<td>23-65</td>
<td>45</td>
</tr>
<tr>
<td>7</td>
<td>40-60</td>
<td>51</td>
<td>41-65</td>
<td>50</td>
</tr>
<tr>
<td>11</td>
<td>21-79</td>
<td>40</td>
<td>25-77</td>
<td>41</td>
</tr>
<tr>
<td>Older Females</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>38-58</td>
<td>47</td>
<td>40-77</td>
<td>43</td>
</tr>
<tr>
<td>9</td>
<td>25-79</td>
<td>45</td>
<td>22-79</td>
<td>49</td>
</tr>
<tr>
<td>10</td>
<td>28-70</td>
<td>45</td>
<td>25-68</td>
<td>43</td>
</tr>
<tr>
<td>Younger Males</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>18-27</td>
<td>20</td>
<td>18-28</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>18-30</td>
<td>20</td>
<td>18-29</td>
<td>21</td>
</tr>
<tr>
<td>8</td>
<td>17-25</td>
<td>20</td>
<td>18-28</td>
<td>20</td>
</tr>
<tr>
<td>Younger Females</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>16-26</td>
<td>19</td>
<td>17-26</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>17-26</td>
<td>19</td>
<td>17-23</td>
<td>19</td>
</tr>
<tr>
<td>12</td>
<td>17-30</td>
<td>20</td>
<td>17-26</td>
<td>20</td>
</tr>
</tbody>
</table>

older and younger groups show some overlap in range of ages but no overlap in median ages. One title in each age-sex group does not allow the subjects free choice of person to fill the title, and the age ranges there are limited by the
ages of specified types of people, i.e., the subjects' parents and girl friends, and the subjects themselves. There can be variation in the ages of people assigned to fill the other two titles in each group. Age ranges and medians were similar for the subjects who were used and the subjects who were rejected for use in the matching study. Here the data will be cited only in regard to the subjects who were used. For the four titles that were free to vary in the two older age-sex groups, the age-sex groups, the age range is from 21 to 79 years; only a few subjects, however, chose people at the extreme ages. For the four titles that were free to vary in the two younger groups, the age range is from 17 to 30 years; only a few subjects chose people at the upper age levels. The median ages of the four sets of people in the two younger groups showed a variation of only one year, from 19 to 20.

The wide range of ages of people selected for the older groups, and the relatively wide range of people selected for the younger groups points up, again, the personal nature of construing. For all the subjects were given the same titles. Some 18 to 20 year old students, for instance, consider a man of 26 to fall in the group "about your own age", while some consider him to be "an older man".
Comparison of the ages of the groups of figures on the two tests shows that the median ages of individual figures within the groups are more variable for the pictured figures than for the Repertory Test figures. On the whole the medians are a little higher for both younger and older age-sex groups on the picture test than they are for the same two groups on the Repertory Test. And also, on the whole, age ranges for the younger groups are less wide on the Repertory Test than on the picture test, and age ranges for the older groups are less wide on the picture test than on the Repertory Test. What age differences exist between figures within the same groups on the two tests appear unlikely to be of much importance for matching accuracy in view of the ability of the judges to match protocols so well when the age-sex groups was varied. (Hypothesis 3c)

The question comes up as to whether the particular lists of titles and sorts that were used influenced strongly the kinds of constructs that were elicited and therefore the relationship between descriptions of figures on the two tests. Since both title and sort lists are short, there is no claim to a complete survey of the construct systems of the subjects. The only relevant data from other research is in terms of the number of different constructs that were elicited. Hunt (thesis) did an investigation of constructs
about a wider variety of people - people over a broader age range and people known in certain specified kinds of situations - than the present experiment used. He also put fewer restrictions on the type of constructs that could be given than did the present experiment. His inspection and comparison of protocols obtained from lists of 20 and 30 titles suggested that the difference in number of constructs obtained from two lists of these lengths was insignificant. His finding that 20 titles were worthwhile points up the fact, of course, that constructs obtained from the present list form only a sample of the total construct repertory of the subjects. It should be remembered, however, that sampling is being done within a restricted area and so a long list is less necessary here than in Hunt's study.

In the same study Hunt varied the number of sorts, giving some subjects up to 80 or 100 sorts, and found that few, if any, new constructs were brought out after the fortieth sort. The number of sorts that are profitable is related to the area of constructs that is being sampled and the length of the title list. Because of the restrictions in area of constructs and length of title list, a long list of sorts appears less necessary here than in Hunt's study. With the restrictions in the present experiment it appears likely that the subjects would stop giving many new constructs
before the fortieth sort had been reached. However, a somewhat longer list of sorts than the one used here would have been desirable if more time had been available for testing sessions and if more pictures had been available so that approximately the same amount of material would still be elicited by the two tests. Observation of the subjects during testing suggests that some, at least, would have given more different constructs if more than 16 sorts had been used, while some had already exhausted their readily available repertory of personality constructs. Only a separate study in which these lists of titles and sorts are used along with longer lists can properly evaluate these lists.

The combinations of titles in the sorts may also influence the kind and number of constructs elicited. The bases on which the combinations were made have already been described in Chapter IV. There it was stated that each sort contained some combination of two figures from the categories of parents, girl friend, self, or ideal figures, plus a disliked person. The combinations were varied so as to allow sorting on the bases of age, sex, whether the people were liked or disliked, or some other relationship such as marriage status or kinship. Some subjects did do part of the sorting in terms of putting together people of similar
age or sex or other relationship, but most of the sub-
jects did most or all of their sorting by dividing the
people according to whether they were liked or disliked,
placing parents, girl friend, and self (when those titles
appeared) with the liked people. This type of sorting
was expected for some subjects, but the extent to which
it occurred was not anticipated. The apparently global
quality of liking and disliking people is consistent with
studies of racial and ethnic prejudice and stereotypes.
Another related study is one in which it was shown that
when errors in marching and calisthenics were made by
a few popular school children, their errors were attribut-
ed to unpopular members of the class. It seems possible
that the subjects liked or disliked the people on some
basis which may have been either unstated or stated in on-
ly a few of their constructs, but that they then made eval-
uations of the people in other areas which might have
little to do, really, with the basis of liking or dislik-
ing the people. The various opinions many subjects ex-
pressed about any one person were too black or too white
to have much likelihood of being arrived at with even re-
relative independence. In terms of personal construct theo-
ry, it may be hypothesized that the majority of these sub-
jects use like-dislike as a superordinate construct which
subsumes a number of other constructs; when the sorts are given in such a way as to contain both liked and disliked people, constellations of dimensions about liked people and disliked people as large groups are brought forth. If such constellations exist and if they are used in regard to figures of varying age and sex, then judges may be expected to be able to match descriptions of people differing in either or both age and sex. Such matching was shown to be possible when the descriptions were about pictured figures, but it was not attempted with the descriptions of real people. Although these subjects did not state whether they liked or disliked the pictured figures, the clinical experience of the experimenter suggests that often subjects do form such judgments, and so it is possible that their descriptions of pictured figures as well as their descriptions of real people were related to like-dislike as a superordinate construct. The experiment was not set up to test use of such a superordinate construct, but its use certainly seems a possibility.

The order in which the sorts occurred is another factor which may influence the number and kinds of constructs elicited. The principal criterion in the original determination of the order of sorts was to reduce as far as
possible the number of times that any title occurred in two consecutive sorts. The best that could be done was the placement of two titles once each in adjacent sorts. Title 5 ("A girl about your own age that you get along well with or would like to know better") occurred in sorts 15 and 16, while title 9 ("An older woman who is nearest your ideal") occurred in sorts 6 and 7. It was believed that if titles did not occur in adjacent sorts, the subjects would be more likely to evaluate the person anew each time his name appeared, whereas if a title appeared in adjacent sorts he might give succeeding responses identical or closely similar to the preceding ones merely because he was set to think in a certain way or because such a response was easier than to evaluate the person again.

It is difficult to evaluate the extent to which set or perseveration of response was important in this experiment. Comments by the subjects suggest that it was important in the case of some individual subjects, that some subjects showed an opposite trend (an aspiration to give as many different adjectives or phrases as possible), and that some subjects gave their responses without much attention either to giving the easiest response or to giving as large a number of different responses as possible. The fact that judges could match protocols obtained from two different
types of tests with a one-week interval between administration of the tests suggests, also, that perseveration is not the only factor involved in the relationship of figure descriptions to each other.

Ways of Improving the Present Experiment

If the experiment were to be done again, several changes would be made. These changes have already been described in various places in connection with other discussions, but they may be summarized here.

1. Preferably, pictures 12M and A4 would not be used because a majority of the misperceptions occurred in response to those pictures. Omission of those pictures would reduce greatly the number of students whom it was necessary to reject as subjects for the matching study. It is conceivable, however, that other published pictures could not be found as substitutes for the two older men and one younger man described in these two pictures. In that case, it might be necessary to retain the two pictures and continue to drop some subjects, for at least three figures per age-sex group appear to be essential for eliciting samples of constructs. As explained before, it is considered desirable, for clinical reasons, to use published pictures rather than unpublished ones drawn es-
especially for this research and then discarded.

2. Although questioning the subjects about perceived similarities of drawings, as distinguished from similarities of interpretations, is not an integral part of the tests of the hypotheses, it can give adjunctive information about the possible bases for the elicitation of constructs about the personalities of the figures. A standard procedure with a more extensive pre-test trial should be worked out.

3. After the subjects have described the figures on the picture test, they should be asked to refer back to the figures when giving the antonyms of the adjectives and phrases used to describe the figures. This procedure would clarify their construct systems somewhat, in case they have two or more antonyms for any descriptive adjectives or phrases. Also, the subjects should give the antonyms orally rather than in written form, so as to reduce the likelihood that mere perseveration or desire to avoid effort might lead to use of a single antonym for a phrase even though the phrase, as originally used on two or more occasions, might have two or more antonyms. (It was noted that some subjects went through their entire protocol, searching for each occurrence of a given adjective and writing the same antonym in each case, and then repeated
the process with another adjective.)

4. Matching of the type done with the picture test protocols for hypothesis 3 should be done with Repertory Test protocols. Also, matching of picture descriptions of one age-sex group and Repertory Test descriptions of another group (a procedure analogous to hypothesis 3c) should be done in order to make a more nearly complete study of areas of generality of constructs.

5. An additional judge should be used, at least for testing hypotheses 3b and 4b, for the distribution of correct matches by Judges M and B differed significantly in regard to those hypotheses.

Relationship of the Present Experiment to Previous Studies.

For the most part, previous matching studies involving so-called projective test protocols have used the Rorschach test and interpretations of the protocols rather than the protocols themselves. With the Rorschach test, the typical procedure has been to have the judges match some sort of case history information with interpretations of the test records (84, 53), or match two sets of interpretations of the same protocols (53). The writer was able to find only two studies in which the method of correct matching was used with picture-test materials. Secord (107) had judges attempt
to match handwriting samples with pairs of Thematic Apperception Test stories, but the experiment was designed as a study of handwriting rather than of the stories, and little analysis was done in terms of the stories. Slutz (111), at psychological meetings in 1941, reported the use of the matching method with picture-test interpretations to investigate "agreement between analyzers; consistency of content; and extent to which the information gained was already familiar to staff members". No other information about the study was available from her, and the slightly fuller report put out by Bell (8:224-225) appears to contain inaccuracies. Harrison (33, 34) is sometimes said (8) to have used the matching method in validation studies of the Thematic Apperception Test; however, he did not use the method of correct matching as used here, but he used a checking of test interpretations with case history materials. Since there are no reported experiments making use of a matching technique to study picture-test materials, not much is to be gained by comparison of present findings with the findings of other matching experiments.

Intrapersonal features common to several written compositions have been studied by the matching method by Allport, Walker, and Lathers (5). They state that subjects
went through their entire protocol, searching for each occurrence of a given adjective and writing the same antonym in each case, and then repeated the process with another adjective.)

4. Matching-of the type done with the picture test protocols for hypothesis 3 should be done with Repertory Test protocols. Also, matching of picture descriptions of one age-sex group and Repertory Test descriptions of another group (a procedure analogous to hypothesis 3c) should be done in order to make a more nearly complete study of areas of generality of constructs.

5. An additional judge should be used, at least for testing hypotheses 3b and 4b, for the distribution of correct matches by Judges M and B differed significantly in regard to those hypotheses.

Relationship of the Present Experiment to Previous Studies

For the most part, previous matching studies involving so-called projective test protocols have used the Rorschach test and interpretations of the protocols rather than the protocols themselves. With the Rorschach test, the typical procedure has been to have the judges match some sort of case history information with interpretations of the test records (84, 58), or match two sets of interpretations of
the same protocols (§3). The writer was able to find only two studies in which the method of correct matching was used with picture-test materials. Secord (107) had judges attempt to match handwriting samples with pairs of Thematic Apperception Test stories, but the experiment was designed as a study of handwriting rather than of the stories, and little analysis was done in terms of the stories. Slutz (111), at psychological meetings in 1941, reported the use of the matching method with picture-test interpretations to investigate "agreement between analyzers; consistency of content; and extent to which the information gained was already familiar to staff members". No other information about the study was available from her, and the slightly fuller report put out by Bell (§224-225) appears to contain inaccuracies. Harrison (33, 34) is sometimes said (§) to have used the matching method in validation studies of the Thematic Apperception Test; however, he did not use the method of correct matching as used here, but he used a checking of test interpretations with case history materials. Since there are no reported experiments making use of a matching technique to study picture-test materials, not much is to be gained by comparison of present findings with the findings of other matching experiments.
Intrapersonal features common to several written compositions have been studied by the matching method by Allport, Walker, and Lathers (5). They state that

"In psychological parlance, we are interested in the discovery of the most highly generalized trait of an individual which seems to be expressed in his written productions. In discussions of composition, this element of the characteristic may be spoken of as individuality.... Individuality in writing is that aspect of quality of writing which is the most reliable means of recognizing a given individual's work and of distinguishing it, in the absence of other identifying features, from the work of another."

In the attempts to match, their judges noted chiefly traits related to the manner of writing, but they also noted, to some extent, the content of the writing. Their study and the present one are similar in that both make use of verbal material, both emphasize the individual, both are interested in general characteristics of some sort. Beyond that, the two studies differ. Allport, Walker, and Lathers were looking for a trait in each individual that is "expressed" in his writing, while the present study emphasizes verbal behavior, not an underlying trait. Their analyses were largely in terms of the manner of writing (e.g., effectiveness, detail, form of composition) and only secondarily in terms of content. The major stress here is on content of the protocols, although other aspects are used secondarily; and the content here is of a particu-
lar type - constructs about the personality or character of people. Despite the differences of material which make direct comparison impossible, the two studies agree in demonstrating that protocols can be matched more accurately than chance alone would suggest, and that the protocols of some subjects are considerably easier to match than the protocols of other subjects. Matching in their study and in part of the present study depends upon the presence of some characteristic(s) extending generally through the records of each individual, plus the presence of discriminably different characteristics in different individuals. But since the theoretical orientations of the two studies differ, there are also differences in the general, yet personal, characteristics that are looked for.

Only a fraction of the studies have used the actual protocols of the subjects rather than the interpretations of the protocols. The studies of secord and of Allport Walker, and Lathers used the original protocols. Where interpretations are used for matching, there is an extra, intermediate step which allows greater opportunity for error. Unless judges match solely on the basis of some such features as identical words, they must do some interpreting, i.e., they must use their own constructs about the protocols. But if interpretations are
given to the judges, the judges must construe another person's constructs about the protocols; they must work one step removed from the actual data. Now when we are interested in the subjects' constructs (or, in some other studies, in the subjects' ideas or methods of expression), why should we make the judges work with another person's constructs? It is suggested that if other investigators wish to use the matching method with projective materials, for instance in further validation studies of the Rorschach technique and the Thematic Apperception Test, they use the original protocols as one of the materials rather than interpretations of the protocols.

Although it has not been done in this report, the subjects' response to the pictures used in this experiment can be worked up into another contribution to the still slowly growing fund of normative data about the Murray cards and other sets of pictures. Eron (24, 25) and Garfield and Eron (28) have given us data in regard to the feeling tones, themes, outcomes, identifications of figures, and perceptual distortions shown in Thematic Apperception Test stories. Rosenzweig and Fleming (95) have listed popular figure identifications, attention to common objects, problems and outcomes, and reaction time, total time, and total wordage for Thematic Apperception Test stories. Wittenborn's norms (130) appear to be his abstractions from the subjects' stories, in contrast
to Rosenzweig and Fleming's classifications on the bases of the subjects' actual words. His list of frequent types of response for ten of the Murray cards is a mixture of traits, activities, figure identifications, evaluations, etc., and it appears less systematic and thorough than the investigations of Eron or even of Rosenzweig and Fleming.

"Picture pull" has been described by Eron (24) in his discussion of norms for the various pictures. He states that "It has been felt that each individual picture has its own stimulus properties which evoke themes, identifications, feelings, etc. which are peculiar to it and which differ from those elicited by other pictures." His lists of frequent themes, feeling tones, etc., demonstrate this. In the present study the accuracy with which judges could match descriptions of figures to the pictures that elicited them is further confirmation of the existence of definite "picture pull". The figure descriptions, if worked up in normative fashion, can clarify the type of content of the "picture pull". The present experiment can actually give more detailed data on figure characterizations than the other studies do, since that is the focus of attention here, and it is only a part of the other studies.

Eron states that the most frequent perceptual distortions among his subjects were sexual misidentifications.
In the present study ten of the misperceptions were in terms of the sex of the figure, but several were in terms of the age of the figure, and two were in terms of both the age and the sex of the figure. Among our subjects sexual misidentifications were only a little more common than age misidentifications. Unless subjects are asked directly to state the ages of the figures, the sexual misidentifications are easier to note than the age misidentifications. It seems possible that Eron missed some of the age misidentifications because of not directly asking subjects to state the ages. Since only part of the pictures were the same in Eron's study and the present one, only a few comparisons can be made as to figures on which sexual misperception occurred. Eron reports misidentification of the reclining figure in picture 12M as a girl or woman by 12% of his nonhospitalized groups and by 13% of his hospitalized groups, while in the present experiment 11% of the subjects made a similar misidentification. He reports identification of the two figures in picture 13MF as either father and daughter or mother and son by 1% of the non-hospitalized and 7% of the hospitalized subjects; in this study a father-daughter identification was made by 6% of the subjects. Other misidentifications of figures in pictures used in both
his study and the present study occurred only rarely. His college students showed more sex confusion than his schizophrenics, although the difference was not significant. The present study confirms the finding that misperceptions occur in a fairly large proportion of college students.

No normative studies comparable to Eron's or Rosenzweig and Fleming's have been reported for the Symonds cards. The present investigation suggests that misidentifications may well be more frequent in response to the Symonds cards than in response to the Murray cards. Differences in the way the figures are drawn appear to be responsible.

Implications of the Experimental Results for Personal Construct Theory and Its Use

It has been demonstrated in this experiment that judges can match descriptions of various real and pictured people so as to put together the descriptions given by the same subjects. As has been pointed out previously, such matching depends upon both the ability to distinguish between the descriptions given by various subjects and the ability to see the similarities in the various descriptions by a single subject. It is held that the two factors involved are the personal quality and the gener-
ality of subjects' constructs. The major problem in this experiment has been the generality of constructs; demonstration of the personal quality is a by-product. In the following pages, however, some implications of both the personal quality and generality of constructs will be discussed.

First a point that should be clarified is the relationship between the constructs elicited by these tests and a person's vocabulary. It is true that the words used by a subject in describing the people are in his vocabulary, but the tests are not merely vocabulary tests. Undoubtedly a subject can understand and use many more words and phrases descriptive of personality than he uses in these tests. What the tests are intended to sample is his repertory of readily available constructs; for if certain constructs are readily available in one situation, then the likelihood is increased that they will be available in other situations, also. Information about an individual's vocabulary can be gained more efficiently by asking him to define a list of words, but only tangential information about his personal construct system can be obtained in that way.

An individual's repertory of constructs about personality and character cannot be assessed adequately merely by a knowledge of the terms he uses to describe
real or pictured people. The terms he uses are held to be merely symbols of constructs which are dimensional in nature, and identical terms can symbolize different constructs for different people. The personal quality of constructs may be illustrated by a list of some of the descriptive terms and their antonyms as used by subjects in this experiment.

<table>
<thead>
<tr>
<th>Descriptive Terms</th>
<th>Antonyms</th>
</tr>
</thead>
<tbody>
<tr>
<td>aggressive</td>
<td>slow, lazy, weak, regressive</td>
</tr>
<tr>
<td>easy going</td>
<td>stubborn, demanding, strict, set in one's own ways, always griping, sadist</td>
</tr>
<tr>
<td>kind</td>
<td>unkind, mean, hard to like, evil, insistent, dominating, greedy, scornful, not pleasant</td>
</tr>
<tr>
<td>quiet</td>
<td>loud, showy, driving, extraverted, sociable, excitable, talkative, gossip, friendly</td>
</tr>
<tr>
<td>stern</td>
<td>weak, friendly</td>
</tr>
<tr>
<td>understanding</td>
<td>mean, narrow-minded, strong minded, unsympathetic, nonbeliever, stern, hard to get along with, inconsiderate, insistent, tends to misinterpret</td>
</tr>
<tr>
<td>independent</td>
<td>dependent, tries hard to please, not self-reliant, interested in other people</td>
</tr>
<tr>
<td>calm</td>
<td>nervous, vigorous, flaring temper, worries, restless</td>
</tr>
<tr>
<td>loving</td>
<td>untrustworthy, disdainful, condescending, hating, nagging</td>
</tr>
</tbody>
</table>
This list points up the fact that if several people use the same word they are not necessarily talking about the same thing. If two people, for instance, describe a person as aggressive, one may mean that he is "not weak" and another that he is "not lazy". At the same time, if those two people hear someone else use the word aggressive, they will interpret his statement in the same light of their own constructs and may well hear two different statements even though both can repeat the speaker's words. Teachers, therapists, and friends in casual conversation may not be communicating what they think they are communicating.

The comparative effectiveness of a speaker's use of the listener's constructs as contrasted with his use of his own constructs is a problem worthy of research. In a child guidance clinic, for example, parents might be given information or recommendations about their child in psychological terms, in terms of the psychologist's own personal constructs, or in terms of the parent's own constructs. It is believed that the latter course will lead to the greatest understanding and ability to carry out recommendations. A therapist working with patients in individual sessions may be better able to help his patients change if he uses their language rather than his own a good part of the time. In a group situation such as a classroom, the speaker obviously cannot use the construct system of each individ-
ual present, and often in teaching situations it would not be desirable even if he could do so. But he can use sufficient illustrative material and can state his major points in enough different ways that the listeners can either understand his constructs or relate his points to their own constructs.

It was noted during this experiment that some terms, as used by a number of people, have only one or two antonyms, while other terms (e.g., "understanding") have at least eight or ten different antonyms. It has been shown in research on the use of rating scales that traits that are rated most consistently are ones that are defined most clearly in terms of overt behavior (Rotter and Wickens). Such a factor is believed to be behind the variability in antonyms for given words. The difference in number of antonyms used with the words "aggressive" and "understanding" may serve as an example. If an individual predicts that someone will be aggressive, he is making a prediction about overt behavior; the behavior can be noted, and various people can compare the terms that they used to predict and describe him. Through such comparisons of observable behavior, various individuals may come to form somewhat similar constructs symbolized by a given term. But if an individual predicts that someone will be understanding, he is making a prediction about behavior that may be covert, or may have less obvious overt cues. Although various individ-
uals may compare the terms that they used to predict him, there is less evidence available for validation or invalidation of their predictions. If there is little to point to, there is a greater chance that each individual will continue to use his verbal symbols in his own way.

Since the personal quality and generality of constructs as they occur on tests was being investigated, we wish to know something of the limitations in the range of the generality. There definitely do appear to be limits to the generality of constructs used on the two tests. Part of the limit may be due to suppression of responses, particularly on the Repertory Test, but the ways of minimizing this have already been discussed. Another limit as noted by Judge B - that the descriptions obtained from the picture test appeared to be "richer" than those obtained from the Repertory Test - appears to be an artifact of the test directions and the organization of the test protocols as given to the judges. In the picture test each figure occurred only once, and the subjects were specifically requested to give three or four different descriptive words or phrases about each figure. In the Repertory Test, on the other hand, each name occurred four times, and the subjects were required to give only one descriptive word or phrase per occurrence of a name. Since there was no requirement that the phrases must be different, some of them were repeated. Also, in the Repertory Test, each de-
scriptive dimension was given about three people at one time, and so there was necessarily considerable similarity of words and phrases applied to several people. As a result there was more frequent repetition of the same terms on the Repertory Test protocols as given to the judges than there was on the picture test protocols. Part of the impression of difference in "richness" may be attributed to this difference in repetition. Even with allowance for this artifact of repetition, subjective impression suggests that a greater variety of terms and somewhat different terms were used on the picture test than on the Repertory Test. The extent of the difference in use of constructs on the two tests seemed to vary from one subject to another. One hypothesis is that some of the subjects may have learned terms about personality through conversation which remained on a largely verbal level; real people and events may never have been included as elements in the constructs symbolized by those terms, with the result that the terms were not available for use when the subjects were asked to describe people they knew well. In other words, the generality of some of the constructs of some people, may be limited to verbal or pictured situations. This puts a restriction on the potential usefulness of picture tests as a means of surveying a person's construct repertory. Obviously, however, all constructs elicited by a picture test are not
limited solely to verbal or pictured situations, for the judges were able to match descriptions of real and pictured people with an accuracy significantly above chance. Further study with other groups of subjects is needed.

Study of the personal quality and generality of individuals' constructs through tests is only a preliminary investigation of the range of generality of constructs. The real question is not whether the generality of personal constructs is wide enough that the same constructs are used in two tests, but whether it is wide enough that the same constructs are used in both tests and non-test situations. After all, there is little point in surveying a person's constructs by means of tests, if those constructs are not applied in other situations. A next step in the investigation might be a study of whether individual patients in a clinic, for example, use the same terms in describing people on a test and in describing people they know while giving social histories or while in therapeutic sessions. It is believed that many of the same constructs will be used in the two situations, although it is not expected that all the constructs that appear in one situation will appear in the other. Since it has been shown that constructs elicited by the pictures used in this experiment are to some extent closely related to the pictures (hypothesis 1), it may be expected that constructs used in an interview will also be related to the situations being discussed. If the interview and the pictures (or other
test materials or questions) deal with situations that are quite similar, they may use some of the same constructs in the two situations, but if the interview and the pictures deal with situations that are quite different, the individuals may or may not use many of the same constructs in the two situations. An attempt to make test pictures representative of common situations (as Murray did in terms of psychoanalytic theory) may be one way of surveying, by tests, constructs that are likely to be used in non-test situations. Another possible solution is to try to use tests to sample the most available constructs (i.e., the ones with the widest generality). For such a purpose it seems likely that the pictures should not show much background or definite action which could specify the situation in which the figures were placed. Most of the Murray and Symonds drawings that were used in this experiment appear to be too explicit for this kind of use.

Personal construct theory holds that an individual's actions will be consistent with his constructs. It may be expected, therefore, that if an individual uses the same construct in two situations, his actions in the two situations will show some relatedness, or, at least, will not be incompatible with each other. The relationship between generality of constructs and compatibility of actions would not be too difficult to test. One possible
procedure would be to have subjects describe a number of people. Certain hypothetical situations might be set up, and the subjects might be asked to predict the behavior of each person in those situations. Their predictions about people described by the same constructs might be compared for similarity or, at least, compatibility.

In the present research two possible areas of generality of constructs were investigated, and it was found that generality did not appear to occur along age or sex lines to any great extent. The subjects used did not differentiate to a perceptible degree between the ages and sex of other people in their application of personality constructs to those people. The question arises as to whether there are other discoverable areas of generality of constructs that are fairly consistently shown by many subjects or whether the areas vary considerably from one subject to another. Now it seems likely that generality of a given construct will occur in regard to people who are all subsumed under the same superordinate construct. The construct "like-dislike" has already been suggested as one such superordinate construct under which generality may occur. On the other hand some psychologists act as if the construct of "authority-non-authority figures" were such a superordinate construct and believe that all people included within the class of authority figures will be
construed in much the same way by any given subject. However, a view consistent with personal construct theory might be that superordinate constructs determining lines of generality, like constructs themselves, are personal and vary from one individual to another. At the time the present research was begun, the writer had only a recent acquaintance with personal construct theory, and she had not thought through this possibility. Under this view it might be expected that some individuals would show generality of constructs within the area of sex, some within the area of age, some within the area of authority figures, some within the area of indulgent people, etc. This variety would mean that for each patient in a clinic, for instance, the area(s) of generality must be determined on an individual basis. Practically, this would be a more difficult task than determination of an area or areas of generality supposedly for whole populations and then application of the findings to individuals on subsequent occasions.

There is not only the possibility that various people may show generality of personality constructs along different lines, i.e., under the regnancy of different superordinate constructs, but also the probability that the degree of generality of major constructs will vary from one person to another. It may be hypothesized that individuals will show a roughly inverse relationship be-
tween a sort of average degree of generality of their constructs about people and the total number of such constructs. It seems likely that individuals who make relatively fine discriminations between other people have more constructs than individuals who believe, for example, that "people are all alike". A person's constructs about people are his perceptions of similarities and differences between people, and if he makes fine discriminations between individuals, he will probably do it, to a great extent, in terms of his past perceptions of similarities and differences between people. All his constructs about people, however, need not be verbalized nor even verbalizable. And every word in a person's vocabulary does not necessarily symbolize a different construct, for a person may have several words to express the same idea. Thus it cannot be said that there is an inverse relationship between degree of generality of constructs and the total vocabulary of individuals.

A factor which may be related to between-person differences in number and degree of generality of personality constructs is restriction of variety of experiences for any reason. Restriction may well mean that the person does not have the opportunity to see contrasts between people that are necessary for the formation of new constructs, and, as a result, his fund of constructs remains limited. Then if he is placed in new situations or meets new people, he must construe them in terms of his small repertory of con-
structs. These new people become additional elements in his old constructs, at least temporarily, and his constructs thus become quite general in their applications. Such broad generality of constructs might come about if experience is restricted through living in a relatively isolated community or a relatively close-knit cultural group, or through a home environment that prevents exploration or contacts with many people.

It seems possible that limitation of number of constructs and broad generality of constructs may also come about when a child is reared by authoritarian parents and teachers whose own constructs show broad generality. Even though a child lives in an area where he meets many kinds of people, if his parents and teachers construe them all in a small number of ways and "hammer home" their own constructions to the child, the child may adopt his parents' limited and overly general constructions.

It may be hypothesized that adjustment difficulties may be associated with either a relatively small repertory of constructs that have a relatively wide generality, or a relatively large repertory of constructs that have a relatively small degrees of generality. A possible example of the former type of relationship was noted by the writer when she was instructor and adviser of a group of 60 first-quarter university freshmen. There were quite a few students from rural or village areas who had fewer than the average
number of ways of construing their new classmates; they seemed to try to construe the classmates in terms of their old repertory of constructs and to become confused and upset in the process. Contacts with students in an undergraduate mental hygiene course, and work in a student counseling service have also given suggestive support to such an hypothesis. Unless an individual is in a close-knit group, it seems likely that if he uses a small number of constructs to predict (construe) the actions of a large number of people, he will often predict incorrectly. If he realizes that his predictions are often invalidated, he may acquire a sense of frustration or inability to cope with problems, or a wish for help in solving problems having to do with other people. If he is not aware that many of his predictions are not valid, and if he acts on his invalid predictions, then other people may call him a problem person. On the other hand, if a person's constructs nearly all have small degrees of generality, he is likely to have difficulty in dealing with new people because he cannot see them as fitting into his already-formed repertory of constructs but must act in terms of specific situations or persons. His behavior is then likely to be unorganized. The hypothesis that so-called maladjusted people are likely to have either above or below average generality of constructs might be tested by
comparison of students who apply for help or are sent to a student counseling service and a random sampling of students who have not sought nor been referred for such help. Generality of their constructs might be investigated by use of two tests similar to the ones used in this experiment. It might be predicted that the matchability of protocols of maladjusted students would show a bimodal distribution, while the matchability of protocols of randomly selected students would be unimodal, with the measures of central tendency falling between the two peaks on the curve for maladjusted students.
CHAPTER IX

Summary and Conclusions

The psychologist working in a clinical situation is often called upon to make predictions about the future behavior of people in a variety of situations. According to the psychology of personal constructs, now being developed by G. A. Kelly (53), such predictions are likely to be more accurate if the psychologist has some knowledge of the person's expectations of other people and the dimensions along which he is free to move. The expectations and dimensions are believed to be shown by the person's constructs. (A construct is defined as a way in which two things are perceived as being alike and yet different from a third thing.) It is held that as a person has participate in or observed events, he has formed constructs about those events, and the constructs have been directed toward optimal anticipation of future events. In construing people, a person predicts what those people will do. His predictions or expectations of the actions of other people, in turn, help to determine his own actions.

A person's constructs about people are believed to be symbolized by the words he uses to describe people, but these words are not the constructs themselves. Since
constructs are believed to be dimensional in nature, a
descriptive term and its antonym provide a better know-
ledge of a construct than does the descriptive term alone.

Both the psychology of personal constructs and the
writer's observations in a clinical setting lead to the
question whether constructs used by a person in one situ-
ation are likely to be used by that person in other situ-
ations. Specifically, the question investigated in the
present research is whether constructs about people elicited
by a picture test and by a form of Kelly's Repertory
Test (52) have sufficient generality to appear in situ-
ations other than the particular test which elicited them.
(The generality of a construct is defined by the number
of elements that the construct subsumes.) The major at-
tention was directed toward the generality of constructs
elicited by the pictured figures, but several related
questions were investigated simultaneously. The experi-
mental variables included particular drawings, age-sex
group membership of pictured figures and real people de-
scribed, individual construction systems, and specific
types of test (here pictures versus names of real people
as stimulus material). The relationship between personal
constructs and the particular pictures that elicited them
was studied because a close relationship implies a limit
to the generality of constructs obtained with pictures.
The relationship between constructs elicited by two different parts of a single (picture) test was studied because if generality of constructs can be shown at all, it should appear here. The relationship between constructs elicited by two different tests was studied because a close relationship implies that constructs obtained on one test are not entirely specific to that test, i.e., they have a wider degree of generality than the one test situation. Two possible areas within which constructs may show generality - age group and sex group - were also studied. The experimental design used made it necessary to investigate the personal quality of constructs at the same time that the generality of constructs was being tested.

The experiment was divided into two parts: the testing of the subjects and the handling of the test data by the method of correct matching. Undergraduate men were given two tests in which they were required to describe the "personality" or "character" of people. Kelly's Repertory Test employed a combination of a forced choice technique and a sorting method to make the subjects compare and describe twelve people (in four age-sex groups) whom they knew well. The picture test consisted of a selection of twelve figures (in four age-sex groups) from the Murray
(80) and Symonds (118) sets of cards; subjects were required to list three or four adjectives or short phrases descriptive of the character or personality of each figure. Both the tests were set up so that the data obtained included both descriptive terms and what the subjects conceived to be the antonyms of the descriptive terms. The protocols or parts of protocols from the two tests were broken up in various ways, and two trained judges were instructed to try to recombine them, matching the parts that belonged together according to certain specifications. The degree of relationship between the materials matched and the significance of the relationship were studied by means of Vernon's modified contingency coefficient (124). The number of correct matches made by the judges was compared with the number to be expected from chance alone by means of critical ratios. The patterns of correct matches made by the judges was compared with the patterns to be expected from chance alone by means of chi-square.

Four main hypotheses were tested, and two of the main hypotheses were divided into sub-hypotheses; thus there were seven hypotheses and sub-hypotheses. Tests of six of the hypotheses and sub-hypotheses were significant at high levels of confidence. The test of the other hypo-
theses was good enough that the results can be considered to have practical as well as statistical significance. The following results were demonstrated for the subjects used.

1. There were both perceptible relationships between certain TAT-type drawings and the descriptions elicited by those drawings, and perceptible distinctions between descriptions elicited by different drawings. The descriptions of some of the pictured figures were more easily identified than the descriptions of other figures. Also, the descriptions of other figures. Also, the descriptions by some subjects appeared to be more closely related to the drawings than did the descriptions by other subjects.

2. There were both perceptible relationships between the descriptions of various pictured figures given by any one subject and perceptible distinctions between the descriptions of the same figures given by different subjects. These relationships and distinctions were shown both when the age-sex group to which the figures belonged was the same for the two sets of materials to be matched and when the age-sex group of the figures in the two sets of materials was different.

3. There were both perceptible relationships between the descriptions of pictured figures and real people given
by any one subject, and distinctions between the descriptions of the same figures and people given by different subjects. Only the case where the age-sex group was the same for the people in the two sets of materials to be matched was studied.

The ability of the judges to match picture descriptions in terms of the age-sex group membership of the figures described appears to be of borderline statistical significance. The borderline character of the statistical tests and the good ability of the judges to match protocols in terms of subjects, even when the age-sex group membership of the two sets of descriptions differed, suggests that these subjects did not show many practically important differentiations between the personalities of pictured figures in terms of the age and sex of the figures.

For all the hypotheses tested the two judges showed better than chance agreement in terms of the number of identical matches. In terms of the pattern of number of correct matches for the various sets of materials, the judges showed no significant difference in tests of five of the hypotheses and sub-hypotheses, while they did differ significantly in tests of two of the hypotheses and sub-hypotheses.

From the results of the experiment the following conclusions appear to be justified.
1. Constructs elicited by pictured figures and by a form of Kelly's Repertory Test do have some generality, at least to another test situation. That is, constructs elicited by these two tests are not wholly specific to the particular test that elicited them. Therefore the possibility is increased that constructs elicited by any one of these tests will be used in other, non-test situations.

2. There is a limitation to the generality of some of the constructs elicited by pictured figures, for it has been shown that for some subjects there are both perceptible relationships between some pictured figures and the constructs elicited by them, and perceptible distinctions between the constructs elicited by different figures. Psychologists should beware of assuming that because a subject describes one person or a small number of people in a certain way, he will necessarily describe many people in the same way. At most, the description of one person or a few people can show what constructs are in a subject's repertory and therefore what ones he may use in other situations.

3. In spite of the above conclusion, constructs elicited by pictured figures apparently are not related solely to the situation in which they are elicited, and there is an important personal quality in the construct systems of
individuals. As an aid in increasing understanding and reducing misunderstanding, it appears advisable for psychologists to determine not only the verbal symbols but also the construct dimensions that are frequently used by their patients.

4. Generality of many of the constructs does not appear to occur in terms of age-sex group, at least for the subjects tested. It is consistent with the psychology of personal constructs to expect that superordinate constructs under whose regnancy generality may occur will vary from one subject to another. Incidental results of this experiment suggest the use by many subjects of the dimension "like-dislike" as such a construct.

5. Occasional misperceptions of the age and/or sex of pictured figures appears to occur in about one-third of college men. Further investigation of their significance is needed.


10. Bellak, L., and Bellak, Sonya S. Children's Apperception Test. N.Y.: C.P.S. Co., P.O. Box 42, Gracie Station, 1949. (Set of 10 cards and manual.)


27. Frank, L.K. Projective methods for the study of personality. J. Psychol., 1939, 8, 389-413.


34. Harrison, R. Studies in the use and validity of the Thematic Apperception Test with mentally disordered patients. III. Validation by the method of "blind analysis". Char. and Pers., 1940, 9, 134-138.


53. Kelly, G.A. The psychological construction of life. (Manuscript)


The Picture Test

In addition to the test directions given to all subjects, certain procedures were set up to be used in case subjects gave specific types of responses. These are described below.

1. If a subject asked, e.g., "Should I tell whether he's cheerful?" he was told again to give whatever words or phrases he thought best described the character or personality of the figure, but not to describe the figure's physical appearance, occupation, or activity in the picture.

2. If a subject gave an answer that appeared to relate to the physical appearance of a figure, the situation in the picture, etc., he was asked whether that response had something to do with the figure's character or personality. If the subject answered affirmatively, his response was recorded. If the subject answered negatively, his response was not recorded, and he was asked to give words descriptive of the figure's personality.

3. If a subject gave vague or ambiguous responses, he was asked to explain what he meant. Such explanations were always requested if subjects used the terms "nice", "good personality", "weak", etc.

4. If a subject gave a response which, while applying
to the personality of a figure, could specifically identify the age-sex group of that figure, he was requested to give a synonym. Requests for synonyms were made when subjects gave the terms "motherly" and "fatherly".

5. If a subject gave an answer that might be relevant to either one particular pictured situation or to character or personality in general, the subject was asked, "Is that what he looks like just here, or is he like that a good deal of the time?" Such an explanation was requested, e.g., if a subject described a frowning figure as "worried".

6. If a subject gave very long answers, he was told, "That's too much for me to write; I don't know shorthand. Out of all that, what am I supposed to write down?"

7. If a subject asked, "Is that enough?" after giving only one or two words about a figure, he was told to "give at least three characteristics for each figure."

8. If a subject asked whether he might skip a figure, on which he was having difficulty, and return to it later, he was told that that was not allowed, but that he should take his time and look at the picture until he thought of a more complete answer.

Repetory Test

In addition to the test directions given to all subjects, certain procedures were set up to be used in case subjects gave specific types of responses. These are de-
scribed below.

1. Sometimes a subject gave his description solely in terms of the type of relationship between the three people, *e.g.*, he might say that two of the people "enjoy each other" and the third person "doesn't enjoy being with the others", or that two of the people "get along well together" and the third person "does not get along with the other two". In that case, the subject was asked to explain how the personalities of the people were related to the response.

2. If a subject gave obviously highly personal reactions to people, he was asked to "Tell what the people themselves are like; don't tell how they affect you." Such a request was made if a subject stated, *e.g.*, that two people were his friends and one was not his friend.

3. If a subject indicated by a comment that he thought he should always sort the liked people together and the disliked person as the contrasting figure, he was told that he might put together any two people that he wished.

4. If a subject gave a response which appeared to relate chiefly to events in which the people participated, he was asked, "What does that imply about their character or personality?" This occurred, *e.g.*, if a subject said that a person had "had many disappointments in life".

5. If a subject started to give more than one grouping per sort, he was told that only one was necessary. If he actually did give more than one grouping, only the first
grouping and its associated verbal responses were used. This procedure was adopted in order to make more nearly equal the amount of information obtained from all subjects.

**Questionnaire for the Repertory Test**

A copy of the questionnaire given to all subjects is shown below.

**Name Questionnaire**

Will you please answer the following questions about the people on your NameaSheet. Be sure that the numbers on this questionnaire correspond with the proper numbers & names on the Name Sheet.

1. Approximate Age ____ College______. Year in College______. Major______.

2. If she is not your mother, what relationship does she have to you? _____________________. Approximate age.

3. Why is the person of your own age that is nearest to your ideal? _______________. Approximate age__.

4. Is (or was) she your girl friend__or your wife___? Approximate age____.

5. Why do you get along well with this girl, or why would you like to know her better? _____________________. Approximate age___________. 
6. Why don’t you like this man? __________________________
   __________________________  Approximate age____

7. If he is not your father, what relationship does he have to you?
   __________________________
   __________________________  Approximate age____

8. Why don’t you like this person? __________________________
   __________________________  Approximate age____

9. Why is she the older woman nearest your ideal?____
   __________________________  Approximate age____

10. Why don’t you like this woman? __________________________
    __________________________  Approximate age____

11. Why is he the older man nearest your ideal?____
    __________________________  Approximate age____

12. Why don’t you like this girl? __________________________
    __________________________  Approximate age____
APPENDIX B

Instructions to the Judges

Introduction

In a clinic setting it is often desirable to know something of the patient's expectations of other people and of the directions in which he is likely to move if he begins to change his own behavior. According to Dr. Kelly's Psychology of Personal Constructs clues as to their expectations and possible directions of movement can be obtained from the words that the patients use to describe events and people. These words are thought of as symbols of constructs (concepts) which represent the patient's abstractions and generalizations from past events and his anticipations of future events. In construing people, a patient predicts what the people will do. His expectations of the actions of other people, in turn, help to determine his own actions toward those people. If he makes similar predictions about two people, he is more likely to act similarly toward them than if he construes the two people differently. Constructs are thought of as being dimensional in nature. The two ends of a personality construct dimension might be represented by an adjective and its antonym. It is believed that in a person's thinking certain dimensions will predominate but that he will place some people at one end and some at the other end of these dimensions. If a person
describes his own behavior or the behavior of others in certain ways, the type of change that he is most likely to show himself or that he is most likely to perceive in others is movement toward the opposite end of the dimension.

General Instructions.

If the clinician is to make use of a patient's constructs, he must have some means of obtaining a sample of those constructs. In this experiment two tests have been used to survey the ideas or constructs of young men (Psychology 401 students) about other people. In both tests they were requested to name personality characteristics of certain people. On one of the tests they were required to describe real people, and on the other test they described pictured people. On both tests the people described varied as to age and sex. Subjects were also asked to state what they considered to be the opposite of each word or phrase used to describe each person. In the parts of protocols which you will be asked to match, the various words or phrases used to describe any one figure are listed together, single spaced. The opposite of each word or phrase is written after it in parentheses. The type of description obtained and the antonyms stated are illustrated by the example given below.

- dominating (indifferent)
- conceited (friendly)
- forward (bashful)
In this illustration the stimulus person was described as being dominating, conceited, and forward. The subject considered that "indifferent" is the antonym of "dominating", "friendly" the antonym of "conceited", and "bashful" the antonym of "forward". "Dominating" is underlined because that was considered to be the most important characteristic of the particular figure described. In some of the sets of descriptions the most important characteristic will be underlined, and in some of the descriptions no such indication will be given. As you go through the materials you will find that the protocol sheets contain descriptions of from one to 12 figures. Adjectives and phrases describing any one figure will be typed together, single spaced. Double spacing between sets of adjectives and phrases will be used to indicate that more than one stimulus person was described. Thus in the example below three different people were described.

| kind       | (unkind) |
| good natured | (grouchy) |
| affectionate | (hatred) |
| unpleasant  | (pleasant) |
| nosey       | (considerate) |
| unfriendly  | (friendly) |
| unpleasant  | (pleasant) |
| cruel       | (kind) |
| thoughtless  | (thoughtful) |

On any one protocol sheet all the descriptions will be by a single subject. The actual protocols given by the subjects have been broken up in various ways. Some of the parts
will be identified by letters and some by numbers. You are asked to put together the parts of protocols that belong together, matching lettered with numbered parts. There will always be the same number of lettered and numbered parts of protocols, and each part can be correctly matched with only one other part. In the large envelopes containing the materials for testing the various hypotheses and subhypotheses you will find specific written directions about the ways in which you are to match parts of protocols.

Please do your best to match all parts of protocols correctly. The success or failure of the study depends heavily on the accuracy of the matchings that you are able to make. If you wish, you will be given further information about the hypotheses and the results after you have completed the task.

**Cues for Use in Matching**

In order that both judges will use the same cues, some of the possible cues will be described below. Of course, you will not be able to use all of them in matching each set of materials. It is up to you to decide which cues apply to any particular set of materials. You may find that other cues occur to you. If so, use them, but please make a note of this.

1. An obvious cue is *similarity or identity of adjec-
tives or phrases used in two parts of a protocol. Sometimes this will be similarity or identity of single adjectives or phrases. More often, probably, sets of two or more adjectives (or phrases) may be repeated in the two parts of any one protocol. These sets may be called descriptive configurations, and usually they will be more important than single adjectives (or phrases).

2. You should attend not only to the figure descriptions themselves, but also to the antonyms. An adjective and its antonym may be considered to be the two poles of a dimension. E.g., a certain subject may use a dimension "kind - unkind", and he will place some people at one end and other people at the other end of the dimension, and perhaps put some people in the middle as being neither very kind nor very unkind. Another subject may use a dimension "understanding - unkind" or "kind - cruel". Subjects may use the same word (e.g., kind) to describe people and yet mean something different by the word; the meaning is made clearer by attention to the antonym listed. The dimension will provide important clues especially when you are given parts of protocols by several subjects, and are asked to put together the parts given by the same subjects.

3. In protocols in which the main characteristics of people are indicated you may find it helpful to pay
attention to the types of characteristics considered to be most important. This is likely to be especially important if, in a set of protocols, several protocols contain similar descriptions, but there are differences in the adjectives that are underlined.

4. You may also find it helpful to form your own abstractions or concepts about the types of figures that are described, and to put together descriptions about which you form similar abstractions. Use of this cue does provide a chance for error through introduction of your own biases, and therefore you are asked to be cautious in using it.

5. The level of abstraction used by different subjects or by a single subject in describing various figures is another cue. That is, are the descriptions in terms of relatively specific acts or are they high level abstractions that could be used to describe many different acts?

6. What might be called vagueness versus definitive quality of the descriptions is another possible cue. In other words, from the lists of adjectives used about people, how much information are you actually given? To what extent might you find important differences among various persons who could be described by a given set of adjectives, or to what extent do the adjectives used delimit the person described?
7. **Vocabulary level** may also be a cue.

In addition to the suggestions about cues given above there are also two cautions.

1. **Do not attempt to use spelling or grammatical form as a cue.** Any errors in spelling are typing errors by the experimenter, since she tried to correct any errors made by the subjects. Grammatical form has also been altered in some cases for the sake of consistency between subjects.

2. As a test of one of the hypotheses you will be asked to match descriptions with the particular pictures that elicited those descriptions. Remember that pictured figures may be described in various ways, and do not consider as the only possible description the set of adjectives that you might use about a figure.

**General Procedures**

1. Read the sections entitled "Introduction", "General Instructions", and "Cues for use in Matching" if you have not already done so. You may refer back to these sections as often as you wish while you are doing the matching.

2. Open one large envelope at a time, read the typed instructions, and do the matching that is requested. When you have finished matching the materials in one large envelope, return the materials to that envelope and then go on to the next large envelope.
Please be careful to keep the materials in the proper envelopes. On the outside of each large envelope you will find a code number. In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>H2</td>
<td>H3b-IV</td>
</tr>
<tr>
<td>H3b-I</td>
<td>H3b-III</td>
</tr>
<tr>
<td>H3b-II</td>
<td>H3b-II</td>
</tr>
<tr>
<td>H3b-III</td>
<td>H3b-I</td>
</tr>
<tr>
<td>H3b-IV</td>
<td>H3c-III</td>
</tr>
<tr>
<td>H3c-I</td>
<td>H3c-II</td>
</tr>
<tr>
<td>H3c-II</td>
<td>H3c-I</td>
</tr>
<tr>
<td>H3c-III</td>
<td>H3a</td>
</tr>
<tr>
<td>H3a</td>
<td>H4b-IV</td>
</tr>
<tr>
<td>H4b-I</td>
<td>H4b-III</td>
</tr>
<tr>
<td>H4b-II</td>
<td>H4b-II</td>
</tr>
<tr>
<td>H4b-III</td>
<td>H4b-I</td>
</tr>
<tr>
<td>H4b-IV</td>
<td>H4a</td>
</tr>
<tr>
<td>H4a</td>
<td>H2</td>
</tr>
<tr>
<td>H1-OM</td>
<td>H1-OF</td>
</tr>
<tr>
<td>H1-YF</td>
<td>H1-YN</td>
</tr>
<tr>
<td>H1-YM</td>
<td>H1-YF</td>
</tr>
<tr>
<td>H1-OF</td>
<td>H1-OM</td>
</tr>
</tbody>
</table>

3. Within each large envelope you will find some small envelopes. Each of the small envelopes contains a single set of materials to be matched. After you have read the instructions in the large envelope, open one small envelope at a time, match the parts of protocols in the envelope, record your matchings on the slip of paper provided in the small envelope, return the protocols and your answer slip to the small envelope, and go on to the next small envelope. Never go back to an envelope that you have previously completed.

4. As you do the matching, will you please make brief notes about any cues that you consider of special im-
importance in helping you to decide which protocols go together. Will you record these notes on the answer slips. You don't need to make notes on all the answer slips, but notes on some of them may be of aid in the interpretation of the judgments. (Please don't write on the protocols themselves.) Notes will be especially important if you use cues other than those listed on pages 4 and 5 of these instructions.

5. As you do the matching, will you, also, please make brief notes about any special difficulties that you experience. Will you record these notes on the answer slips, unless you believe that your statements apply to many sets of materials. In that case, will you use a separate sheet of paper and state the extent to which you believe that the difficulty exists. These notes may facilitate interpretation of the results especially if the hypotheses are not supported.

Matching Procedure: Hypothesis 1-OM

In the large envelope you will find 3 pictures, each containing one figure commonly said to be a male of an age generation older than that of the subjects. A slip of paper fastened to each picture designates which figure is said to be the older male and gives the identification number assigned to that picture for this matching task. In each of the small envelopes you will find 3 slips of paper, each identified by a letter. Each paper contains a de-
scription elicited by one of the 3 numbered figures. The descriptions on all 3 papers were given by a single subject. Will you try to match lettered papers with numbered pictures so that each description is paired with the figure that elicited it. For example, if you believed that the figure identified by the number 10 elicited the description lettered F, you would write F after the number 10 on the answer slip. In each of the small envelopes there is one description given in response to each pictured figure, and so you should match each paper with one of the pictures, and match each paper with only one of the pictures. Remember that pictured figures may be described in various ways, and do not consider as the only possible description the set of adjectives that you might use about a figure.

**Order of Matching:** Notice that the small envelopes are all identified by the code letters "Hl-OM" with another number following this. The last number identifies each individual envelope. Use that last number as a guide to the order in which you are to open the small envelopes. (Open only one at a time, do the matching required, record your choices, and return the materials to that envelope before opening another.) In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>47</td>
<td>2</td>
</tr>
<tr>
<td>46</td>
<td>3</td>
</tr>
<tr>
<td>44</td>
<td>5</td>
</tr>
<tr>
<td>43</td>
<td>7</td>
</tr>
</tbody>
</table>
Matching Procedure: Hypothesis 1-OF

In the large envelope you will find 3 pictures, each containing one figure commonly said to be a female of an age generation older than that of the subjects. A slip of paper fastened to each picture designates which figure is said to be the older female and gives the identification number assigned to that picture for this matching task. In each of the small envelopes you will find 3 slips of paper, each identified by a letter. Each paper contains a description elicited by one of the 3 numbered figures. The
The descriptions on all 3 papers were given by a single subject. Will you try to match lettered papers with numbered pictures so that each description is paired with the figure that elicited it. For example, if you believed that the figure identified by the number 10 elicited the description lettered F, you would write F after the number 10 on the answer slip. In each of the small envelopes there is one description given in response to each pictured figure, and so you should match each paper with one of the pictures, and match each paper with only one of the pictures. Remember that pictured figures may be described in various ways, and do not consider as the only possible description the set of adjectives that you might use about a figure.

Order of Matching: Notice that the small envelopes are all identified by the code letters "HL-OF" with another number following this. The last number identifies each individual envelope. Use that last number as a guide to the order in which you are to open the small envelopes. (Open only one at a time, do the matching required, record your choices, and return the materials to that envelope before opening another.) In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>49</td>
</tr>
<tr>
<td>2</td>
<td>47</td>
</tr>
<tr>
<td>3</td>
<td>46</td>
</tr>
<tr>
<td>5</td>
<td>44</td>
</tr>
<tr>
<td>7</td>
<td>43</td>
</tr>
<tr>
<td>9</td>
<td>42</td>
</tr>
<tr>
<td>10</td>
<td>41</td>
</tr>
<tr>
<td>12</td>
<td>39</td>
</tr>
</tbody>
</table>
Matching Procedure: Hypothesis 1-YM

In the large envelope you will find 3 pictures, each containing one figure commonly said to be a young male. A slip of paper fastened to each picture designates which figure is said to be the young male and gives the identification number assigned to that picture for this matching task. In each of the small envelopes you will find 3 slips of paper, each identified by a letter. Each paper contains a description elicited by one of the 3 numbered figures. The descriptions on all 3 papers were given by a single subject. Will you try to match lettered papers with numbered pictures so that each description is paired with the figure that
elicited it. For example, if you believed that the figure identified by the number 10 elicited the description lettered F, you would write F after the number 10 on the answer slip. In each of the small envelopes there is one description given in response to each pictured figure, and so you should match each paper with one of the pictures, and match each paper with only one of the pictures. Remember that pictured figures may be described in various ways, and do not consider as the only possible description the set of adjectives that you might use about a figure.

Order of Matching: Notice that the small envelopes are all identified by the code letters "H1-YM" with another number following this. The last number identifies each individual envelope. Use that last number as a guide to the order in which you are to open the small envelopes. (Open only one at a time, do the matching required, record your choices, and return the materials to that envelope before opening another.) In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th></th>
<th>JE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>49</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>47</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>46</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>44</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>43</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>42</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>41</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>39</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>36</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td>34</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>33</td>
</tr>
<tr>
<td>21</td>
<td></td>
<td>31</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td>30</td>
</tr>
</tbody>
</table>
Matching Procedure: Hypothesis 1-YF

In the large envelope you will find 3 pictures, each containing one figure commonly said to be a young female. A slip of paper fastened to each picture designates which figure is said to be the young female and gives the identification number assigned to that picture for this matching task. In each of the small envelopes you will find 3 slips of paper, each identified by a letter. Each paper contains a description elicited by one of the 3 numbered figures. The descriptions on all 3 papers were given by a single subject. Will you try to match lettered papers with numbered pictures so that each description is paired with the figure that elicited it. For example, if you believed that the figure identi-
fied by the number 10 elicited the description lettered F, you would write F after the number 10 on the answer slip. In each of the small envelopes there is one description given in response to each pictured figure, and so you should match each paper with one of the pictures, and match each paper with only one of the pictures. Remember that pictured figures may be described in various ways, and do not consider as the only possible description the set of adjectives that you might use about a figure.

**Order of Matching:** Notice that the small envelopes are all identified by the code letters "Hl-YF" with another number following this. The last number identifies each individual envelope. Use that last number as a guide to the order in which you are to open the small envelopes. (Open only one at a time, do the matching required, record your choices, and return the materials to that envelope before opening another.) In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th>JF</th>
</tr>
</thead>
<tbody>
<tr>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>47</td>
<td>2</td>
</tr>
<tr>
<td>46</td>
<td>3</td>
</tr>
<tr>
<td>44</td>
<td>5</td>
</tr>
<tr>
<td>43</td>
<td>7</td>
</tr>
<tr>
<td>42</td>
<td>9</td>
</tr>
<tr>
<td>41</td>
<td>10</td>
</tr>
<tr>
<td>39</td>
<td>12</td>
</tr>
<tr>
<td>38</td>
<td>13</td>
</tr>
<tr>
<td>36</td>
<td>14</td>
</tr>
</tbody>
</table>
Matching Procedure: Hypothesis 2

In each small envelope you will find 8 slips of paper. There are 4 numbered papers, each of which contains a description of a single person. There are also 4 lettered papers, each of which contains descriptions of 2 people. All the descriptions in any one small envelope were given by a single subject. One numbered paper and one lettered paper contain descriptions of males in approximately the same age generation as the subject, one numbered and one lettered paper contain descriptions of females in approximately the same age generation as the sub-
ject, one numbered paper and one lettered paper contain descriptions of males who are older than the subject, and one numbered and one lettered paper contain descriptions of females who are older than the subject. Will you try to match numbered and lettered papers so that descriptions about people of the same sex and roughly similar ages are put together. In many cases you may believe that certain numbered and lettered papers go together but cannot identify the age-sex group to which the descriptions refer. Identification of the groups is not required. For example, if you believed that the paper lettered Z and the paper numbered 25 both contained descriptions about young males, you would put those 2 papers together. In recording your matching you would put those 2 papers together. In recording your matching you would merely put the letter Z after the number 25 on the answer slip; you would not be required to write "young males" on the answer slip. In each of the groups, and so you should match each paper with some other paper, and match each paper with only one other paper.

Order of Matching: Notice that the small envelopes are all identified by the code letters "H2" with another number following this. The last number identifies each individual envelope. Use that last number as a guide to the order in which you are to open the small envelopes. (Open
only one at a time, do the matching required, record your choices, and return the materials to that envelope before opening another.) In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>49</td>
</tr>
<tr>
<td>2</td>
<td>47</td>
</tr>
<tr>
<td>3</td>
<td>46</td>
</tr>
<tr>
<td>4</td>
<td>45</td>
</tr>
<tr>
<td>5</td>
<td>44</td>
</tr>
<tr>
<td>6</td>
<td>43</td>
</tr>
<tr>
<td>7</td>
<td>42</td>
</tr>
<tr>
<td>8</td>
<td>41</td>
</tr>
<tr>
<td>9</td>
<td>40</td>
</tr>
<tr>
<td>10</td>
<td>39</td>
</tr>
<tr>
<td>11</td>
<td>38</td>
</tr>
<tr>
<td>12</td>
<td>37</td>
</tr>
<tr>
<td>13</td>
<td>36</td>
</tr>
<tr>
<td>14</td>
<td>35</td>
</tr>
<tr>
<td>15</td>
<td>34</td>
</tr>
<tr>
<td>16</td>
<td>33</td>
</tr>
<tr>
<td>17</td>
<td>32</td>
</tr>
<tr>
<td>18</td>
<td>31</td>
</tr>
<tr>
<td>19</td>
<td>30</td>
</tr>
<tr>
<td>20</td>
<td>29</td>
</tr>
<tr>
<td>21</td>
<td>28</td>
</tr>
<tr>
<td>22</td>
<td>27</td>
</tr>
<tr>
<td>23</td>
<td>26</td>
</tr>
<tr>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>25</td>
<td>24</td>
</tr>
<tr>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>27</td>
<td>22</td>
</tr>
<tr>
<td>28</td>
<td>21</td>
</tr>
<tr>
<td>29</td>
<td>20</td>
</tr>
<tr>
<td>30</td>
<td>19</td>
</tr>
<tr>
<td>31</td>
<td>18</td>
</tr>
<tr>
<td>32</td>
<td>17</td>
</tr>
<tr>
<td>33</td>
<td>16</td>
</tr>
<tr>
<td>34</td>
<td>15</td>
</tr>
<tr>
<td>35</td>
<td>14</td>
</tr>
<tr>
<td>36</td>
<td>13</td>
</tr>
<tr>
<td>37</td>
<td>12</td>
</tr>
<tr>
<td>38</td>
<td>11</td>
</tr>
<tr>
<td>39</td>
<td>10</td>
</tr>
<tr>
<td>40</td>
<td>9</td>
</tr>
<tr>
<td>41</td>
<td>8</td>
</tr>
<tr>
<td>42</td>
<td>7</td>
</tr>
<tr>
<td>43</td>
<td>6</td>
</tr>
<tr>
<td>44</td>
<td>5</td>
</tr>
<tr>
<td>45</td>
<td>4</td>
</tr>
<tr>
<td>46</td>
<td>3</td>
</tr>
<tr>
<td>47</td>
<td>2</td>
</tr>
<tr>
<td>48</td>
<td>1</td>
</tr>
</tbody>
</table>

**Matching Procedure: Hypothesis 3a**

In each small envelope you will find 8 slips of paper. There are 4 numbered papers, each of which contains de-
scripions of 4 people; all the people described on any one paper belong to different age-sex groups - i.e., each person described differs from the other 3 people described on that paper in sex or in age generation or in both sex and age. There are also 4 lettered papers, each of which contains descriptions of 8 people; on each lettered paper 2 of the people are males of approximately the same age generation as the subjects, 2 are females of approximately the same age generation as the subjects, 2 are males of an older generation, and 2 are females of an older generation. On the lettered papers the 2 descriptions of people in any one age-sex group are placed next to each other. The first person described on the numbered papers and the first 2 people described on the lettered papers belong to the same age-sex group, the second person described on the numbered papers and the third and fourth people described on the lettered papers belong to the same age-sex group, etc. The numbered descriptions were given by 4 different subjects, and the lettered descriptions were given by the same four subjects. (The descriptions on any one paper were all given by a single subject.) Will you try to match numbered and lettered papers so that descriptions given by the same subjects are put together. For example, if you believed that the paper numbered 13 and the paper lettered R contained descriptions by the same subject, and that the paper lettered T and the paper numbered 15 contained descrip-
tions by the same subject, you would write R after 13 and T after 15 on the answer slip. In each of the small envelopes there is one pair of papers for each of the 4 subjects, and so you should match each paper with some other paper, and match each paper with only one other paper.

Order of Matching: Notice that the small envelopes are all identified by the code letters "K3a" with another number following this. The last number identifies each individual envelope. Use that last number as a guide to the order in which you are to open the small envelopes. (Open only one at a time, do the matching required, record your choices, and return the materials to that envelope before opening another.) In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

Matching Procedure: Hypothesis 3b-I

(Note: The matching for hypothesis 3b was broken up into four parts. In part I, the descriptions matched were all about older males. In part II, the descriptions matched were all about older females. In part III, the descriptions matched were all about younger males. In part IV, the descriptions matched were all about younger females.)
In each envelope you will find 8 slips of paper. There are 4 numbered papers, each of which contains a description of a single person. There are also 4 lettered papers, each of which contains descriptions of 2 people. All the people described on both the numbered and lettered papers are of the same sex and roughly similar age group. (They might be either males or females, and they might be in approximately the same age generation as the subjects or in an older generation.) The numbered descriptions were given by 4 different subjects, and the lettered descriptions were given by the same 4 subjects. (In the case of lettered papers, the 2 descriptions on any one paper were given by a single subject.) Will you try to match numbered and lettered papers so that descriptions given by the same subjects are put together. For example, if you believed that the paper numbered 13 and the paper lettered R contained descriptions by the same subject, and that the paper numbered 15 and the paper lettered T contained descriptions by the same subject, you would write R after 13 and T after 15 on the answer slip. In each of the small envelopes there is one pair of papers for each of the 4 subjects, and so you should match each paper with some other paper, and match each paper with only one other paper. **Order of Matching:** Notice that the small envelopes are all identified by the code letters "H3b-I" with another number following this. The last number identifies each individual
envelope. Use that last number as a guide to the order in which you are to open the small envelopes. (Open only one at a time, do the matching required, record your choices, and return the materials to that envelope before opening another.) In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>

Matching Procedures: Hypotheses 3b-II, 3b-III, and 3b-IV
(For the other three parts of the matching for hypothesis 3b, the judges were given the same instructions as for hypothesis 3b-I. Only the order of matching sets within the parts was varied.) For hypothesis 3b-II, the order was:

<table>
<thead>
<tr>
<th>JK</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

For hypothesis 3b-III, the order was:

<table>
<thead>
<tr>
<th>JM</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>
Matching Procedure: Hypothesis 3c-I

(Note: The matching for hypothesis 3c was broken up into three parts. In part I, the descriptions matched were about younger males and younger females. In part II, the descriptions matched were about younger males and older females. In part III, the descriptions matched were about younger males and older males.)

In each small envelope you will find 8 slips of paper. There are 4 numbered papers, each of which contains descriptions of 3 people; all 3 people are of the same sex and roughly similar age group (either approximately the same generation as the subjects or an older generation.) There are also 4 lettered papers, each of which contains descriptions of 3 people; all 3 people are of the same sex and
roughly similar age group (either approximately the same generation as the subjects or an older generation). The age-sex group of the people described on the lettered papers differs from that of the people described on the numbered papers; the group may differ in either age or sex or in both age and sex. For example, all the numbered descriptions might be about young females, and all the lettered descriptions about older females. The numbered descriptions were given by 4 different subjects, and the lettered descriptions were given by the same 4 subjects (All the descriptions on any one paper were given by a single subject.) Will you try to match numbered and lettered papers so that descriptions given by the same subjects are put together. For example, if you believed that the paper numbered 13 and the paper lettered R contained descriptions by the same subject, and that the paper numbered 15 and the paper lettered T contained descriptions by the same subject, you would write R after 13 and T after 15 on the answer slip. In each of the small envelopes there is one pair of papers for each of the 4 subjects, and so you should match each paper with some other paper, and match each paper with only one other paper.

Order of Matching: Notice that the small envelopes are all identified by the code letters "H3c-I" with another number following this. The last number identifies each individual envelope. Use that last number as a guide to the order in which you are to open the small envelopes.
(Open only one at a time, do the matching required, record your choices, and return the materials to that envelope before opening another.) In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th>JE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>

Matching Procedures: Hypotheses 3c-II and 3c-III

(For the other two parts of the matching for hypothesis 3c the judges were given the same instructions as for hypothesis 3c-I. Only the order of matching sets within the parts was varied). For hypothesis 3c-II, the order was:

<table>
<thead>
<tr>
<th>JM</th>
<th>JE</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
</tr>
</tbody>
</table>

For hypothesis 3c-III, the order was:

<table>
<thead>
<tr>
<th>JM</th>
<th>JE</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
Matching Procedure: Hypothesis 4a

In each small envelope you will find 8 slips of paper. There are 4 numbered papers, each of which contains descriptions of 12 people. There are also 4 lettered papers, each of which contains descriptions of 12 other people. On each paper (either numbered or lettered) 3 descriptions are about males of approximately the same age generation as the subjects, 3 are about females of approximately the same age generation as the subjects, 3 are about males of an older generation, and 3 are about females of an older generation. The three descriptions about people in any one age-sex group are placed next to each other and are set off from the descriptions of people of a different age-sex group by horizontal lines on the protocol sheets. Because there are 4 age-sex groups, the descriptions on each paper are divided into 4 sets. The order in which the age-sex groups occur on the numbered sheets is the same as the order in which they occur on the lettered sheets, but the specific people described on the numbered and lettered sheets are different. For example, the first group of 3 descriptions on both the numbered and lettered sheets might be about older males, the second group of 3 descriptions about young females, etc. The numbered descriptions were given by
the same 4 subjects. (The descriptions on any one paper were all given by a single subject.) Will you try to match numbered and lettered papers so that the descriptions given by the same subjects are put together. For example, if you believed that the paper numbered 13 and the paper lettered r contained descriptions by the same subject, and that the paper numbered 15 and the paper lettered T contained descriptions by the same subject, you would write R after 13 and T after 15 on the answer slip. In each of the small envelopes there is one pair of papers for each of the 4 subjects, and so you should match each paper with some other paper, and match each paper with only one other paper. Order of Matching: Notice that the small envelopes are all identified by the code letters "H4a" with another number following this. The last number identifies each individual envelope. Use that last number as a guide to the order in which you are to open the small envelopes. (Open only one at a time, do the matching required, record your choices, and return the materials to that envelope before opening another. In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>
Matching Procedure: Hypothesis 4b-I

(Note: The matching for hypothesis 4b was broken up into four parts. In part I, the descriptions matched were all about younger females. In part II, the descriptions matched were all about younger males. In part III, the descriptions matched were all about older females. In part IV, the descriptions matched were all about older males.)

In each small envelope you will find 8 slips of paper. There are 4 numbered papers, each of which contains descriptions of 3 people. There are also 4 lettered papers, each of which contains descriptions of 3 other people. All the people described on both the numbered and lettered papers are of the same sex and roughly similar age group. (They might be either males or females, and they might be in approximately the same age generation as the subjects or in an older generation.) The numbered descriptions were given by 4 different subjects, and the lettered descriptions were given by the same 4 subjects. (All the descriptions on any one paper were given by a single subject.) Will you try to match numbered and lettered papers so that descriptions given by the same subjects are put together. For example, if you believed that the paper numbered 13 and the paper lettered R contained descriptions by the same subject, and that the paper numbered 15 and the paper lettered T contained descriptions by the same subject, you would write R after 13 and T after 15 on the answer slip. In each of the small envelopes there is one pair of papers for each
of the 4 subjects, and so you should match each paper with some other paper, and match each paper with only one other paper.

Order of Matching: Notice that the small envelopes are all identified by the code letters "H4b-I" with another number following this. The last number identifies each individual envelope. Use that last number as a guide to the order in which you are to open the small envelopes. (Open only one at a time, do the matching required, record your choices, and return the materials to that envelope before opening another.) In doing the matching please follow the order given below under your initials.

<table>
<thead>
<tr>
<th>JM</th>
<th></th>
<th>JB</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Matching Procedures: Hypotheses 4b-II, 4b-III, 4b-IV

(For the other three parts of the matching for hypothesis 4b, the judges were given the same instructions as for hypothesis 4b-I. Only the order of matching sets within the parts was varied.)

For hypothesis 4b-II, the order was:
For hypothesis 4b-III, the order was:

<table>
<thead>
<tr>
<th>JM</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

For hypothesis 4b-IV, the order was:

<table>
<thead>
<tr>
<th>JM</th>
<th>JB</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>
I, Rachel Jane Hamilton, was born in Marysville, Kansas, on November 12, 1919. I received my secondary school education in the public schools of Highland Park and New Brunswick, New Jersey. I attended the New Jersey College for Women, Rutgers University, from which I received the degree Bachelor of Arts in 1941. I was granted a fellowship at Smith College, and received the degree Master of Arts in psychology in 1942 from that college. During the school year 1942-43 I attended the Ohio State University, beginning work on a doctoral degree in clinical psychology. At that time I held a University Scholarship. From 1943 to 1949 I worked as psychologist at the E.P. Bradley Home in East Providence, Rhode Island, and at the Veterans Administration Mental Hygiene Clinic in Philadelphia. In the fall of 1949 I returned to the Ohio State University to continue study toward a doctoral degree, and remained at this university until the end of the summer quarter 1952. During the period at this university I served as teaching assistant and assistant instructor in psychology, and also worked on a temporary or part-time basis as psychologist at the Harding Sanitarium and at the Children's Mental Health Center in Columbus. Since October 1, 1952, I have been employed as chief psychologist at the Guidance Center for Dayton and Montgomery County in Dayton, Ohio. While in Dayton I have been registered for off-campus research in order to complete the requirements for the degree Doctor of Philosophy.