PERSONALITY CHARACTERISTICS ASSOCIATED WITH ACCURACY IN PERSON PERCEPTION: A MULTIVARIATE APPROACH

DISSERTATION

Presented to the Graduate Council of the North Texas State University in Partial Fulfillment of the Requirements

For the Degree of
DOCTOR OF PHILOSOPHY

by

Allyn Kay Hale, B.A., M.S.
Denton, Texas
May, 1979
Doctor of Philosophy (Counseling and Student Services), May, 1979, 155 pp., 18 tables, references, 102 titles.

The purpose of this study was to investigate the personality characteristics associated with accuracy in person perception. The particular personality variables chosen (experience, intelligence, cognitive complexity, psychological mindedness, emotional stability, introversion, and empathic style) for investigation were those previously identified by Gordon Allport as being related to person perception. Both types of person perception (sensitivity to individual differences and sensitivity to group differences) were measured.

The subjects consisted of 115 volunteers (60 females and 55 males). At least 10 students from each of the following groups made up the sample: psychology, education, counselor education, science, business, and fine arts. Other groups containing at least 10 individuals were the following: experienced clinicians (all had PhD's), secretaries, psychiatric patients, and others.

Each subject completed the following: Personal Data Sheet, Culture Fair IQ Test, Eysenck Personality Questionnaire, and three scales (Psychological Mindedness, Flexibility, and
Hogan's Empathy Scale) from the California Psychological Inventory. Sensitivity to individual differences was assessed by having the subjects identify emotions from 36 photographs of facial expressions. Answering MMPI items like males and females, 16-19 years old, and males and females over 70, was the measure of sensitivity to group differences.

The major findings of the study were the following:

1. The most capable judges of individual differences, in comparison with the least capable judges, were significantly more likely to have the following psychometric characteristics: higher IQ, more extraverted, more psychologically minded, more cognitively complex/flexible, and more empathic. Experientially, the most competent judges were likely to be older, older in age at first job, and have fewer siblings. In contrast to the least able judges, the most competent ones had resided in more different cities, traveled more, experienced more formal education, and were more likely employed in the helping professions.

2. The best judges of group differences, in comparison with the worst judges, were more likely female, older in age at first job, and higher scorers on Neuroticism.

3. Significant multiple linear regression equations were formed between the predictor variables and the two judgment tasks, suggesting that accuracy in judging others consists of many skills, not just one.
4. Those who were in the helping professions were better judges than those not in the helping professions on sensitivity to individual differences, but not sensitivity to group differences.

5. For those in the helping professions, years of experience in the helping professions were not related to success on either of the two judgment tasks.
ACKNOWLEDGEMENT

The items used in the present study from the California Psychological Inventory were reproduced by special permission from Consulting Psychologists Press, Inc., and Harrison Gough, author of the California Psychological Inventory.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>iv</td>
</tr>
<tr>
<td><strong>Dissertation</strong></td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td></td>
</tr>
<tr>
<td>Review of the Literature</td>
<td></td>
</tr>
<tr>
<td>Hypotheses</td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>34</td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
</tr>
<tr>
<td>Instruments</td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td></td>
</tr>
<tr>
<td>Collection and Treatment of Data</td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>62</td>
</tr>
<tr>
<td>Summary of Results</td>
<td></td>
</tr>
<tr>
<td>Discussion</td>
<td>101</td>
</tr>
<tr>
<td>Limitations</td>
<td></td>
</tr>
<tr>
<td>Conclusions</td>
<td></td>
</tr>
<tr>
<td>Recommendations</td>
<td></td>
</tr>
<tr>
<td>Appendices</td>
<td>133</td>
</tr>
<tr>
<td>References</td>
<td>144</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>1.</td>
<td>Correlational Matrix of Experience Variables</td>
</tr>
<tr>
<td>2.</td>
<td>Rotated Factor Matrix</td>
</tr>
<tr>
<td>3.</td>
<td>Pearson Product-Moment Correlations between Psychometric Variables and the Two Judgment Tasks</td>
</tr>
<tr>
<td>4.</td>
<td>Point Biserial Correlations between Dichotomous Variables and the Two Judgment Tasks</td>
</tr>
<tr>
<td>5.</td>
<td>Means, Standard Deviations, and t-Values for the Two Judgment Tasks: Males (M) vs. Females (F)</td>
</tr>
<tr>
<td>6.</td>
<td>Pearson Product-Moment Correlations between Measures of Experience and the Two Judgment Tasks</td>
</tr>
<tr>
<td>7.</td>
<td>Means and Standard Deviations for the Various Groups on the Stereotype Accuracy Measure</td>
</tr>
<tr>
<td>8.</td>
<td>Means and Standard Deviations for the Various Groups on the Emotion Recognition Test</td>
</tr>
<tr>
<td>9.</td>
<td>Means, Standard Deviations, and t-Values between the Helpers and Non-helpers on the Two Judgment Tasks</td>
</tr>
<tr>
<td>10.</td>
<td>Means, Standard Deviations, and t-Values between the Helpers and Psychiatric Patients on the Two Judgment Tasks</td>
</tr>
<tr>
<td>11.</td>
<td>Means, Standard Deviations, and t-Values between the Helpers and Fine Arts' Group on the Two Judgment Tasks</td>
</tr>
<tr>
<td>12.</td>
<td>Means, Standard Deviations, and t-Values between the Helpers and Scientists' Group on the Two Judgment Tasks</td>
</tr>
<tr>
<td>13.</td>
<td>Summary Statistics for Multiple Correlation Criterion: Stereotype Accuracy Measure for All Subjects</td>
</tr>
</tbody>
</table>
List of Tables--Continued

14. Summary Statistics for Multiple Correlation Criterion: Emotion Recognition Test for All Subjects .............................................. 90

15. Summary Statistics for Multiple Correlation Criterion: Emotion Recognition Test for Female Subjects Only ................................................. 92

16. Summary Statistics for Multiple Correlation Criterion: Emotion Recognition Test for All Male Subjects ................................................. 94

17. Means, Standard Deviations, and t-Values between the 19 Best Judges and the 18 Worst Judges on the Emotion Recognition Test (ERT) ................................................. 95

18. Means, Standard Deviations, and t-Values between the 17 Best and 18 Worst Judges on the Stereotype Accuracy Measure ................................................. 97
Statement of the Problem and Review of the Literature

Statement of the Problem

The ability to assess, evaluate, and understand others accurately is the *sine qua non* of human interaction. The term most often used to describe this process is person perception, although many related terms such as social perception, person cognition, social cognition, interpersonal perception, *la connaissance d'autrui*, clinical judgment, and empathy have also been used. While few would quibble over the importance of being able to recognize and empathize with the feelings and thoughts of others, little is really known about the personality correlates associated with this phenomenon.

The scientific study of person perception began with Charles Darwin's monumental work *The Expression of the Emotion in Man and Animal* (1872). In this century the most frequently cited theoretician in the field of personality variables and person perception has been Gordon Allport. He identified what he considered to be the most relevant personality variables associated with the ability to judge others: experience, intelligence, cognitive complexity, psychological mindedness, emotional stability, introversion, aesthetic interests, self-insight, and an empathic style.
The problem selected for investigation in this study was to explicate the bivariate and multivariate relationships between selected personality variables identified by Gordon Allport and measures of person perception.

Various researchers over the years have conducted research using only a few of the above-mentioned variables and some measure of empathy. Research by Bronfenbrenner, Harding, and Gallwey (1958) indicated that judging others involved numerous skills, not just one. Despite this finding almost all research in this area has emphasized a bivariate research design. The present study, while examining the bivariate relationships between numerous personality variables and measures of person perception, also has the explicit purpose of studying this phenomenon in multivariate fashion.

While most scientific investigations in this area have limited themselves to one judgment task, Bronfenbrenner et al. (1958) and Cronbach (1955) noted that two distinctly different person perception tasks exist. One involves sensitivity to group differences; the other, sensitivity to individual differences. The former is called stereotype accuracy, sensitivity to the generalized other, or sensitivity to group differences. These terms are employed interchangeably throughout this paper. Knowing what the "typical response" of middle America would be to a particular social issue would be an example of stereotype accuracy.
A second type of skill in person perception focuses on individuals. Several synonymous terms are employed to describe this phenomenon: sensitivity to individual differences, interpersonal sensitivity, differential accuracy, empathy, and clinical judgment. These terms are used interchangeably throughout this work. Basically, this process refers to the ability to recognize the ways in which one person may differ from another. A brief example might best help distinguish the two judgment tasks. A store manager may be an excellent judge of knowing what merchandise will sell well in a certain neighborhood (stereotype accuracy), but a poor judge (differential accuracy) of which customers will buy what products. One of the purposes of the present study is to investigate the relationships between both types of person perception and their relationship to personality variables.

The field of person perception constitutes one of the most important areas for those interested in understanding the counseling process. Much literature (Bergin, 1971; Fiedler, 1950, 1951; Luborsky, Chandler, Auerbach, Cohen, & Bachrach, 1971; Rogers, Gendlin, Kiesler & Truax, 1967; Truax & Carkhuff, 1967) has accumulated which has emphasized the tremendous importance of counselor understanding of client as being one of the curative agents of counseling. Consistently, writers in the field of psychopathology (Costello, 1970; Millon, 1969; Sullivan, 1953) have stressed the role of misperception of others' feelings, inaccurate recognition
of stimuli, and poor ability to understand others as characteristic of those with difficulties in interpersonal relationships, faulty life styles, and severe mental disorders.

Carkhuff (1969) has documented not only the lack of effective people in the helping professions, but he has forcefully argued that lack of empathic ability is one of the main reasons that too few profit from professional help. With greater understanding of those characteristics associated with empathy, those responsible for recruiting future helpers could either make better selections for professional school training, or they could train their students to be more effective.

Review of the Literature

Given that knowledge of good judges of others is most important, the question naturally is raised of why more is not known. Paucity of research is not the problem. Rather, numerous methodological problems exist. These problems have been so serious that such a noted authority as Cronbach (1955) has seriously questioned the results and interpretations of almost all the pre-1955 studies.

One of his major criticisms was that "being a good judge of others" was not just one construct. Rather, he saw it as being made up of four separate components. The two most important are stereotype accuracy and differential accuracy. The two lesser components are elevation
and differential elevation. For a complete discussion of the methodological problems in person perception research, the reader is referred to the seminal articles on this topic (Cline, 1964; Cronbach, 1955; Gage & Cronbach, 1955; Taguiri, 1969). A sample of the most important methodological problems are discussed below.

The $D^2$ (difference) statistic was employed by Taft (1950) as one measure of interpersonal accuracy. Each subject was required to rate all the other subjects, himself, and how he thought the others would evaluate him on a five point continuum on six different personality traits. The subjects were all participants in a weekend assessment study. The criterion of accuracy of judgment was the consensus of the experts and the peer ratings from others. Judgmental accuracy was measured by taking the difference between the predicted score and the obtained score. This score was subtracted and squared. The individual with the lowest score was the best judge. The error in this design was that it rewarded those who took a very cautious approach and consistently gave the middle score of three. Choosing such a strategy the individual was never more than two from the predicted score ($5 - 3 = 2$, or $3 - 1 = 2$). The result of this methodological defect was that accuracy was much more a function of response style than any intrinsic measure of accuracy in person perception.
Problems associated with social desirability have also plagued researchers in this area. Social desirability refers to the problem of individuals giving answers to questions in a style that emphasizes the socially acceptable. Those who give only socially desirable answers on a personality questionnaire are assumed to be "faking good." Stelmachers and McHugh (1964) had judges answer over 200 items from the Minnesota Multiphasic Personality Inventory (MMPI) and 16PF (Personal Factors). The judges were instructed to answer those items separately for each of the following four target persons: (1) a normal female teenage college sophomore, (2) an elderly female psychiatric patient who had received the diagnosis of "conversion reaction," (3) an adult male psychiatric patient with a diagnosis of "depression reaction," and (4) an unemployed delinquent teenaged boy. In one comparison condition, the subjects had extensive information about the four target persons. Yet, had the subjects assumed each of the above four target persons would have answered the personality inventory items in the socially desirable direction, they would have been significantly more accurate \( (p < .001) \) for three of the four subjects than if they had utilized the extensive information available to them!

In a similar study Wolfe, Young, and Bryant (1977) had subjects answer 182 MMPI items like they thought two filmed interviewees would answer them. Like Stelmachers and McHugh, they found that if the subjects had merely answered all the
items in the socially desirable direction, they would have produced higher scores than actually trying to guess the two subjects' responses. Hence, a social desirability response style could considerably influence scores on any judgment task.

Additionally, one of the most frequently encountered response style difficulties in person perception research is the assumed similarity problem. This means that a judge assumes that a target person would answer items the same way that he would. In the previously cited study by Stelmachers and McHugh, the judges would have been more accurate for two of the four target persons if they merely projected their own answers onto the subjects, rather than trying to accurately guess them. Accuracy in judgment then would have been a function of similarity of judge and target person, rather than any skill in interpersonal sensitivity.

Along with the assumed similarity problem, the use of stereotypes has been one of the most serious methodological problems in doing research on sensitivity to individual differences. This is the reason for inclusion in the study of a measure of stereotype accuracy. Research (Cline, 1964; Sines, 1959; Stelmachers & McHugh, 1964) has indicated that the use of stereotypes is often more precise than having extensive information on each subject. In the Sines' study, judges (clinical psychologists) described 30 patients by a 97-item Q-sort. (A Q-sort consists of a group of cards, in
this case 97 cards, which have one descriptive statement on each card. The judges are asked to sort the cards in several different stacks ranging from "most characteristic" to "least characteristic." In addition, they completed a Q-sort on the "typical mental patient" (stereotype description). The five psychologists were told to make a Q-sort after receiving each of the following pieces of input information: MMPI, Rorschach, Biographical Data Sheet, and Interview. As part of a counter-balancing design, the information was presented in different order to each of the diagnosing clinicians. After assessment, each of the patients was given 10 psychotherapy sessions. At the conclusion of those sessions, the therapist completed the 97-item Q-sort, which became the criterion of accuracy of judgment. While it is gratifying that the clinicians were more accurate than the stereotypical description in 21 of the 30 cases, nine (almost one-third) of the clients were more accurately described by the stereotypical description of the typical mental patient, despite the extensive data available to the clinicians! The point is clear: stereotype components contribute heavily to accurate person perception.

In the same vein, one of the biggest issues in clinical psychology has been the clinical vs. statistical prediction argument, which was begun by Meehl (1954). In brief, Meehl compared psychologists' prognostications on various issues clinicians are frequently asked to predict (grade point average, recidivism rates among prisoners, and diagnosis) with some
type of actuarial description (usually involving multiple linear regression equations). Meehl challenged psychologists to show him an example of psychologists functioning more accurately than an actuarial description. Since Meehl's challenge, over 50 studies have been published on this topic. No study has had the clinician doing better than the actuarial description (Goldberg, 1970; Meehl, 1965; Wiggins, 1973).

The point is that a "good" stereotypical description (also called base rate) often provides a more accurate description of someone in a particular setting than the best clinical psychologist's prediction. One can see that someone who was extensively familiar with various stereotypes (base rates of a particular setting) would make a better judge of others than someone relying on a more intuitive approach. One can appreciate how this could confound research on individual differences in sensitivity. Cronbach (1955) expressed the thought that success in judging others would much more reside in the judge (knowledge of stereotypes) than any particular characteristic of the target person to be judged.

Homogeneity of subjects has also caused difficulty in interpretation of findings. For example, Taft's well-known study (1950) had 40 graduate students at the University of California as subjects. Sharp's recent study (1974) found no relationship between IQ and measures of person perception, yet all his subjects were graduate students in counseling!
Very likely with such a restricted range on IQ, finding significant correlations with it would have been almost impossible. The present study has attempted to correct this methodological problem.

One method of increasing the heterogeneity of the subjects is to select subjects from a wide variety of occupations. Holland (1966) has presented a typology of occupations. He believes that most individuals can be categorized as one of the following six types: Realistic, Intellectual, Social, Conventional, Enterprising, and Artistic. As these various "types" are different from each other, it was decided that one method of obtaining a more diverse sample would be to select subjects from as many of the Holland types as possible. The result was that at least one group (ten individuals) of each of those six vocational types, except Realistic, was formed for the present study.

Given all the methodological problems that exist in person perception research, no one should be surprised that numerous problems are encountered in trying to compare various studies in person perception. More often than not conflicting findings exist in the literature. As apparent from the previous discussion of methodological difficulties, those problems account for differences in many of the studies. Various schema have been developed to examine all the factors
involved in person perception research. After reviewing the literature, there appear to be the following major divisions involved in doing research in this area:

1. Stimulus information used in the study
2. Nature of the judgment task
3. Criterion information
4. Scoring procedures
5. Characteristics of the judge (personality, intelligence, and emotional state)

This division was based on a distillation of similar ideas given by both Cline (1964) and Taguiri (1969). The focus of the present study is on the last category: characteristics of the judge.

To appreciate the complexities involved in doing research in this area, a brief review is made of each of the above-mentioned characteristics involved in research in the area. Since the focus of the present study involves the personality characteristics of the judge, an extensive review of the literature is given to that topic.

The stimulus information in person perception research varies enormously from study to study. Photographs, motion pictures, tape recordings, written material (biographical data), test scores (IQ's, MMPI's, and Rorschach protocols), and face-to-face interactions have all been used as stimuli
in this area (Cline, 1964). One can readily see that these different cues would lend themselves to certain predictions more easily than others.

The types of judging instruments also differ considerably. For example, in one research design, subjects were asked to label facial expressions as either happy or sad (Levy, Orr, & Rosenzweig, 1960). More complicated designs have subjects write psychological reports based on the information available to them. Other types of judgment instruments are the following: predicting responses to personality tests, predicting scores on tests (IQ and personality), making a Q-sort description, matching personality descriptions with test data, and matching emotional expressions with photographs of facial expressions.

Criterion information changes from study to study. Criterion information refers to what is the correct answer in a person perception task. For example, experts (experienced psychologists) may decide on a target person's diagnosis (criterion), then ask the judges (graduate students) to make their guess as to the correct diagnosis. Self-ratings of the target person often serve as the criterion. For example, the individual's responses to the MMPI are the criterion to be predicted. Associates of the target person (roommates, peers, and therapists) often provide information about the target person.
Numerous different scoring procedures exist. Number of correct answers to a multiple-choice test or a true-false format is a quite commonly met scoring procedure. Making the correct diagnosis is also one of the most frequently encountered judgment tasks. The $D^2$ (a difference score) statistic used to be employed with rating scales. As noted earlier, certain methodological problems exist with this method.

Personality characteristics of the perceiver constitute the major focus of this study. The following personality variables, each noted by Gordon Allport, were selected for investigation in this study: experience, intelligence, cognitive complexity, psychological mindedness, emotional stability, introversion, and an empathic style. The literature for each, as it relates to person perception, is discussed below.

Experience, according to Allport (1961), is the first requirement of a good judge. Basically he believed that a wide range of experience with human nature, not just attainment of chronological adulthood, was necessary before one could be a good judge. "Without a broad and long continued acquaintance with all manner of people, the elementary basis for logical inference is lacking" (p. 507). Despite the intuitively logical appeal of experience, obtaining quantifiable measures of it, other than age, are difficult to find. For this reason, very little research has been conducted in this area.
Developmental trends (Dymond, Hughes, & Raabe, 1952; Gates, 1923; Hall, Rosenthal, Archer, DiMatteo, & Rogers, 1977) from childhood to early adulthood have been observed. However, with adults, little is really known. Reviewing the early research in the area, Taft (1955) cautioned that most studies using adults have had a very narrow range of ages; therefore, definitive statements about age and person perception are unavailable. The recent research by Hall et al. (1977) has suggested that a linear trend goes from childhood to early adulthood, with a leveling off occurring somewhere between 20 and 30 years of age. The present study had a broad age range (16 - 52 years). Hence, greater knowledge of the relationship between age and person perception should be gained.

Sex differences in person perception research have not been consistently observed. However, when they have been found, they have typically favored women (Hall et al., 1977; Hoffman, 1977; Taguiri, 1969; Wolfe, 1974). This is consistent with the prevailing sex role stereotype that females are more empathic than males. Hoffman made the distinction between two kinds of empathy: (1) recognition of affect, and (2) vicarious affective arousal. The latter he considered a better measure of empathy. He cited nine studies which measured vicarious affective arousal. In all cases the females scored higher than the males. The 12 studies examining recognition of affect showed no sex differences.
While the number of studies given by Hoffman in support of his views was impressive, the subjects in all his studies except one were small children!

Maccoby and Jacklin (1974) reviewed 29 studies of empathy and sex differences. In most studies there were no sex differences. However, the females were superior in six studies; males, three studies. Again, most of the studies involved children, not adults.

Why sex differences in empathy might exist is open to much conjecture. Freud (1925/1961) believed that females were more guided in their social interactions by affect than by reality considerations. This was based on his belief that females do not identify with the parent of the same sex as quickly as males. This was due to anatomical differences which did not necessitate their resolution of the Oedipal situation as rapidly as males.

Hoffman believed that males are socialized for an active mastery set toward life, whereas females are more oriented for an affective orientation; therefore, females should be better at empathy. While researchers have examined sex differences in person perception research, the final word has not been written on this topic. In addition, no known research exists comparing the sexes on a stereotype accuracy measure. Compared to most of the earlier research cited, the present study has a more adult-aged sample.
Although the relationship between sex differences and person perception is unclear, the association between IQ and person perception is much more straightforward. Intelligence and person perception have been researched rather extensively. Allport's view (1937), although dated now, still best summarizes this association:

Experimental studies have found repeatedly that some relationship exists between superior intelligence and the ability to judge others... even within a high and narrow range of intelligence... Understanding people is largely a matter of perceiving relations between past and present activities, between cause and effect, and intelligence is the ability to perceive just such relations as these. (p. 514)

Most research (Davitz, 1964; Gates, 1923; Kanner, 1931; Taft, 1955; Taguiri, 1969; Weisgerber, 1956; Wolfe, 1974) has agreed with Allport. No studies have found negative correlates; however, some research has found no relationship (Levy, Orr, & Rosenzweig, 1960; Sharp, 1974). Taguiri (1969) concluded that intelligence is the most consistent of the personality variables in its relationship to measures of person perception.

The one other cognitive measure in this study was cognitive complexity. Allport (1961) believed it to be
one of the most meaningful constructs in person perception:

As a rule people cannot comprehend others who are more complex and subtle than they. The single track mind has little feeling for the conflicts of a versatile mind. People who prefer simplicity of design and have no taste for the complex in their aesthetic judgments are not as good judges as those with a more complex cognitive style and taste. (p. 508)

The preeminent investigator in the area of cognitive complexity and judgment of others has been Bieri. George Kelly's Rep Test has been his main measure of cognitive complexity. Bieri (1955) found a significant positive relationship between cognitive complexity and the ability to predict behavior in common social situations. Interestingly enough he found that cognitive complexity was much more positively related ($r = .35$) to accurate perception of differences between judge and others than between similarities between judge and others ($r = .02$). Leventhal (1957) found a similar trend, although not statistically significant. Hence, cognitive complexity appears to be especially pertinent in noting differences between people, but not as important in judging similarities between people.

Another interesting facet of the cognitive complexity research has been the "felt confidence" of the judge after he has made his judgment. Tripoldi and Bieri (1964) varied
the amount of complexity of the judgment tasks and then acquired confidence measures. The observation was made that highly cognitively complex judges were consistently less confident of their judgments than low cognitively complex judges for all stimulus conditions except those involving incongruent, i.e., complex, information (Bieri, Atkins, Briar, Leaman, Miller, & Tripoldi, 1966).

The antithesis of cognitive complexity is cognitive rigidity. This is one of the characteristics of the authoritarian personality. In one very creative study, Scodel and Mussen (1953) studied individuals scoring very high and very low on a measure of authoritarianism. Allport (1961) summarized this study:

Scodel and Mussen tested college students on a scale designed to measure authoritarianism. They then paired each high authoritarian with another student of the same age and sex who stood low on the same scale.

For twenty minutes these students conversed informally about radio, television, or the movies. Through this casual conversation each formed an impression of the other. None knew the real purpose of the experiment.

After conversing, each student was taken to a separate room and given a questionnaire to fill out as he thought his partner would fill it out.
The results show that the high authoritarians projected their own attitudes; that is, they thought their interlocutor would answer the questions in an authoritarian manner (although, of course, their partners were all low on this scale). By contrast the nonauthoritarian students estimated the attitudes of their partners more correctly. They not only perceived them as authoritarians, which they were, but also estimated more correctly other traits in their natures ...

The authoritarian [cognitively simple] ... believes that his own mental type is universal. He assumes that others are similar to himself, and rigidly projects his convictions upon them. (p. 510)

The measure of cognitive complexity in this study was the Flexibility (Fx) scale of the California Psychological Inventory (CPI). While cognitive complexity is not identical with cognitive flexibility, the two are quite similar constructs. The cognitively complex individual, by definition, is cognitively flexible (Bieri, Atkins, Briar, Leaman, Miller, & Tripoldi, 1966). Hence, because of this relationship between the two variables, the construct cognitive complexity will appear throughout this paper as cognitive complexity/flexibility. Also, in this study the Fx scale of the CPI is interpreted as cognitive complexity. At the
time the study was begun, no personality questionnaire measure of cognitive complexity was known; however, the just-published Jackson Personality Inventory contains such a measure (Jackson, 1977).

Fx has been employed in one previous study (Hjelle, 1969) of person perception. The subjects completed the CPI twice. First, they answered the items, then they completed those same items as they believed their roommate would. The results of that experiment were in the predicted direction, i.e., the most accurate judges of their roommates had higher Fx scores, although this difference fell short of statistical significance. Interestingly enough, the best judges scored significantly higher on the Tolerance scale of the CPI. The latter correlates positively with measures of flexibility, and negatively with rigidity and authoritarianism.

Just as cognitive complexity/flexibility refers to an intellectual style, so does psychological mindedness. Psychological mindedness refers to a type of thinking that is concerned with subjective states--such as inner states, purposes, feelings, fantasies, wishes, and meanings. One who is psychologically minded is an individual who very much believes in psychological determinism. Attributing one’s behavior to the position of the stars or Lady Luck may be correct, but it does not represent psychological mindedness.
In a study of psychological mindedness, Hjelle (1969) had college students complete six scales from the CPI twice, once for self-description and once for prediction of his roommate's responses. He compared good judges of others with poor judges. The results indicated that the two groups differed significantly on a measure of psychological mindedness, with the high scorers being better judges. The measure employed was the Psychological Mindedness scale (Py) from the CPI. The same scale was used in the present study.

Vingoe and Antonoff (1968) studied Py and person perception using a research design very similar to the one used by Hjelle. College students rated both individuals from their dormitory and themselves. Contrary to the Hjelle study, no significant differences were found on Py between the good and poor judges of others.

Wolitzky and Reuben (1974) also studied the relationship between psychological mindedness and sensitivity to individual differences; however, unlike the previous two studies, their measure of psychological mindedness was the degree of psychological mindedness in the subjects' TAT stories. The judgment task consisted of making personality interpretations based on two target persons' TAT stories. The results supported the view that those scoring high on a measure of psychological mindedness had greater precision in making judgments about personality. In conclusion, psychological mindedness appears to be related to person perception.
That the emotional state of the perceiver is an important variable in person perception almost goes without saying. The paranoid knows others are angry and plan to hurt him; the insecure sees everyone as a threat. Much face validity exists concerning the relationship between the emotional state of the perceiver and the accuracy of his perceptions of others; yet hard data has not been plentiful, nor has it been clearcut.

In one well-known study Murray (1933) found that young girls saw photographed faces as more malicious after playing a game of "murder." Schiffenbauer (1974) varied the emotional state of his subjects before they judged emotions from photographs. His results indicated that the emotional state of his judges significantly influenced their judgments.

Psychometric measurements of psychopathology have generally supported the view that those individuals scoring high on measurements of neuroticism and psychoticism, even in normal populations, do worse on measures of interpersonal perception. This seems to be true no matter what instrument is employed. Results using the MMPI (Bergin & Jasper, 1969; Bergin & Solomon, 1970), the Eysenck Personality Inventory (Hekmat, Khajavi, & Mehryar, 1974; Vingoe & Antonoff, 1968), the CPI (Hjelle, 1969; Vingoe & Antonoff, 1968), and Lanyon's Psychological Screening Inventory (Hekmat, Khajavi, & Mehryar, 1975) have all supported this view. Interestingly,
a measure of social nonconformity (a measure to assess antisocial tendencies from Lanyon's Psychological Screening Inventory) was found (Hekmat et al., 1975) to be positively correlated with a measure of empathy, although it did not reach statistical significance. When one considers the general syndrome of psychopathy, these results are not so surprising.

Perceptual anomalies in schizophrenia have long been considered part of the syndrome. Hence, findings (Dougherty, Bartlett, & Izard, 1974; Winkelmayer, Exline, Gottheil, & Paredes, 1978) that schizophrenics did worse than normals in recognizing affect from photographs is not too surprising.

Despite the fact that research on perception and schizophrenia has been plentiful and the results unequivocal, research in the area of extraversion-introversion and person perception has been surprisingly sparse and equivocal. Allport (1961) believed that a healthy detachment (introversion) from others was necessary to be a good judge of others. He said the following about extraversion and its relationship to person perception:

Poor judges are often overly social, excessively affiliative, dependent, or nurturant. We all know the warmhearted soul who feels so much sympathy, pity, love, admiration for others that he (or she) cannot take an impartial view of their failings. And we also know good judges...who participate
heartily enough at times, but also withdraw and consider events and people with an almost grim objectivity. (p. 509)

Vingoe and Antonoff (1968) investigated the relationship between extraversion and person perception. The latter was measured by having the subjects estimate the rating they would receive from dorm mates on six different personality ratings. Extraversion was assessed by the Eysenck Personality Inventory (EPI). The results indicated that good judges were significantly more introverted than poor judges. The findings of both Vernon (1933) and Taft (1950) suggested that introversion was more associated with being a good judge.

Hekmat, Khajavi, and Mehryar (1974, 1975) related extraversion-introversion to Hogan's Empathy Scale, a measurement which has been shown to be highly correlated with numerous ratings of good judges of others. Results from both studies revealed significant positive correlations between extraversion and empathy. Similarly, Hall, Rosenthal, Archer, DiMatteo, and Rogers (1977) found that extraversion, measured by the Myers-Briggs Type Indicator, was associated with nonverbal sensitivity.

Brown and Hendrick (1971) evaluated whether introverts or extraverts were more easily perceived by others. Both introverts and extraverts found that extraverts were more accurately perceived. This finding is in agreement with Allport's
view (1961) that the best judges are introverts, but the best judges are often the least well-known to others.

Most of the personality variables so far discussed are static variables, e.g., IQ. Concerning the actual process of understanding others, Allport stressed the importance of an empathic style, although other methods (inference, analogy) were mentioned by him. As empathy is the broadest of these processes, it was the construct utilized to assess the procedure for knowing others.

Numerous definitions of empathy exist. Most of these fall into two broad categories (Hoffman, 1977). One involves recognition of affect; the other, vicarious affective response to another's feelings. The present study focused on the former.

Different attempts (Cassell, 1963; Chapin, 1942; Dymond, 1949; Hogan, 1969) have been made to develop an objective measure of empathy. Hogan's Empathy Scale, one of the most valid and reliable measures of empathy, was employed in the present study. Hogan demonstrated that his scale correlated .65 with Q-sort descriptions of the "mentally healthy man" and .58 with ratings of social acuity. Hence, measures of empathy were correlated not only with social acuity but with mental health ratings.

Hogan also correlated results from his scale with numerous personality measures (California Psychological Inventory, MMPI, Allport-Vernon Study of Values, Eysenck Personality Inventory, and the Adjective Check List).
The highest correlations Hogan found were with scales that measured social and interpersonal adequacy. Hence, the high scorers on Hogan's Empathy Scale function effectively both socially and interpersonally. The description given by Hogan of the high scoring individuals were based on the Adjective Check List.

The 10 adjectives that Hogan found to be most characteristic (had the highest correlations) of high scorers on his measure were the following:

- pleasant
- charming
- friendly
- dreamy
- cheerful
- sociable
- sentimental
- imaginative
- discreet
- tactful

Hogan portrayed the high scorers on his scale as "likable and friendly, possessing considerable charm, poise, and tact. [They are] outgoing, warm, and very much at ease in the interpersonal situation--the term urbane might even be appropriate" (p. 315).

The 10 adjectives that were most characteristic of the low scorers on Hogan's Empathy Scale were the following:

- cruel
- cold
- quarrelsome
- hostile
- bitter
- unemotional
- unkind
- hard-hearted
- argumentative
- opinionated
Obviously the low scorers on this scale are not easy people to relate with socially, or interpersonally.

Formal education and working experience in the helping professions (psychology, counseling, psychiatry, social work) would seem to be logically related to being a good judge of others. Our culture has certified helpers as "experts" in making important life decisions. It is assumed that by virtue of their training and experience, those individuals are better able to sort through complex amounts of information about an individual and arrive at decisions about him that are superior to the layman. The following question naturally arises: Is there empirical evidence that those in the helping professions make better judgments than those in other occupations?

Sarbin, Taft, and Bailey (1960) reviewed 14 different studies comparing those with an education in psychology with individuals in other fields. The results were not encouraging. Three studies found psychologists worse than laymen; six indicated no significant differences; and only five suggested that psychologists were more accurate judges. Wiggins (1973) summarized more recent research in this area: "There is little empirical evidence that justifies the granting of 'expert' status [as an expert judge] to the clinician on the basis of his training, experience, or information-processing ability" (p. 131).
To give the reader a flavor of the research that has been done, one classic study in this area is described. Goldberg (1959) selected a judgment task that is familiar to most clinicians: the diagnosis of organic brain damage from the Bender-Gestalt test. The latter consists of nine geometric designs that are presented to the client one at a time. The client is instructed to draw each one of them on a sheet of paper. Because visual-motor skills are required to successfully complete this task, the supposition is made that this test is a good indicator of an organic brain syndrome. The experimenter gave each judge 30 different Bender-Gestalt tests to evaluate as either "organic" or "nonorganic." Psychiatric patients drew 15 of the Bender-Gestalts and "organics" drew the other 15.

The judges consisted of three groups. There were four PhD psychologists with four to nine years experience using the Bender-Gestalt. The next group consisted of 10 psychology trainees with MA degrees and one to four years experience. The last group consisted of eight hospital secretaries with no experience with the test. The secretaries were given the test to familiarize them with it, and they were told its purpose.

The results were most illuminating. The mean correct percent accuracy scores for the groups were the following: PhD psychologists, 65%; MA psychologists, 70%; secretaries, 67%. No statistically significant differences were noted.
among the groups. Goldberg noted that only 12 of the 22 judges exceeded chance. One must remember that 50% accuracy would be expected by chance. Goldberg found that the highest percentage of those 12 who exceeded chance was found among the secretaries (62%), the MA trainees (60%), and lastly, the PhD's (25%)!

Given that the icon of "expert status" has not been supported by the data, one might question the sacrosanct status enjoyed by "experienced clinicians." Among those in the helping professions, the expectancy that experienced helpers would be better judges than those who are inexperienced does not seem that unreasonable. Sarbin, Taft, and Bailey (1960) reviewed the early research in this area. They compared those students who had little training in psychology with those who either had graduate training or who were graduates with considerable training. Two studies showed superiority of the better trained judges; two demonstrated that the more experienced judges were worse; and six studies found no differences.

The best known of the early studies was completed by Kelly and Fiske (1951). They compared the judgmental accuracy of beginning and advanced graduate students. The judgment task consisted of predicting responses of psychiatric patients to an inventory. These patients had previously been given psychological tests by these same graduate students. The results indicated no significant difference between the two levels of graduate students.
Goldberg (1968) examined the post-1960 literature on this topic. He rather discouragingly concluded that the amount of professional training and experience of the judges did not relate to judgmental accuracy. He noted that this trend appeared in most research studies cited in the literature.

If indeed the clinical-psychological method that stresses study of the whole individual is not that accurate in making judgments, the question could be asked about the feasibility of employing statistical procedures to address the problem. As noted earlier, this is somewhat tangential to Meehl's clinical vs. statistical prediction controversy. The multivariate research design is the most frequently mentioned statistical procedure that is compared with clinical judgments; therefore, examination of how this method has been employed in person perception research is relevant.

Multivariate experimental methods, despite a long history beginning with Galton and Pearson, have not been successfully applied to research in person perception. This is indeed surprising as the various authorities (Bronfenbrenner, Harding, & Gallwey, 1958; Taguiri, 1969) in the field strongly suggested that person perception involves a multiplicity of skills.

The multivariate approach, unlike the bivariate approach that emphasizes one stimulus and one response, focuses more on the wholeness of the individual. That is, much like the
clinical-psychological method that stresses the study of the whole person, the multivariate research design is oriented towards the study of the entire person. This is accomplished by taking many measurements from the same person. The difference between the two approaches is that the multivariate approach is quantitative and follows explicit calculations of laws and general conclusions. Hence, the clinical approach is really the multivariate approach without the benefit of the apparatus, i.e., statistics (Cattell, 1965).

Only two known studies (Sharp, 1974; Zlatchin, 1973) have attempted to study person perception in multivariate fashion. Both had serious methodological weaknesses. Sharp's study had too few subjects \((n = 40)\), his subjects were highly homogeneous (all graduate students in counseling), and he had no measure of stereotype accuracy. Zlatchin's study was marred by a test of recognition of facial affects that had "disastrously low reliabilities." Hence, research using a multivariate design is much needed.

**Hypotheses**

The following hypotheses were made on the basis of previous research findings:

1. There will be a significant positive correlation between the two types of person perception (stereotype accuracy and differential accuracy).
2. There will be no significant correlation between sex and stereotype accuracy.

3. There will be no significant correlation between sex and differential accuracy.

4. There will be a significant positive correlation between intelligence and stereotype accuracy.

5. There will be a significant positive correlation between intelligence and differential accuracy.

6. There will be a significant positive correlation between cognitive complexity and stereotype accuracy.

7. There will be a significant positive correlation between cognitive complexity and differential accuracy.

8. There will be a significant positive correlation between Psychological Mindedness and stereotype accuracy.

9. There will be a significant positive correlation between Psychological Mindedness and differential accuracy.

10. There will be a significant negative correlation between Neuroticism and stereotype accuracy.

11. There will be a significant negative correlation between Neuroticism and differential accuracy.

12. There will be a significant negative correlation between Psychoticism and stereotype accuracy.

13. There will be a significant negative correlation between Psychoticism and differential accuracy.

14. There will be a significant negative correlation between Extraversion and stereotype accuracy.
15. There will be a significant negative correlation between Extraversion and differential accuracy.

16. There will be a significant positive correlation between an empathic style and stereotype accuracy.

17. There will be a significant positive correlation between an empathic style and differential accuracy.

18. For those in the helping professions, there will be a significant positive correlation between years of experience in the helping professions and stereotype accuracy.

19. For those in the helping professions, there will be a significant positive correlation between years of experience in the helping professions and differential accuracy.

20. The various groups in the study will differ significantly on a measure of stereotype accuracy.

21. The various groups in the study will differ significantly on a measure of differential accuracy.

22. People in the helping professions will be significantly better judges than those in the fine arts on stereotype accuracy.

23. People in the helping professions will be significantly better judges than those in the fine arts on differential accuracy.

24. People in the helping professions will be significantly better judges than those in the sciences on stereotype accuracy.
25. People in the helping professions will be significantly better judges than those in the sciences on differential accuracy.

26. There will be a significant multiple correlation between a combination of predictor variables (intelligence, extraversion, emotional stability, psychological mindedness, cognitive complexity, empathic style, and experience measures) and the criterion variable of stereotype accuracy.

27. There will be a significant multiple correlation between a combination of predictor variables (intelligence, extraversion, emotional stability, psychological mindedness, cognitive complexity, empathic style, and experience measures) and the criterion variable of differential accuracy.

**Method**

**Subjects**

A narrative describing this research was filed with the Human Subjects Review Board at North Texas State University. The Board subsequently approved it. A copy of that letter is enclosed in Appendix A. All subjects were volunteers. The only inducement to the subjects was a promise of a summary of the overall results at the completion of the research.

The subjects of this study consisted of 115 subjects. Approximately 70% of the subjects were students, faculty, and employees of North Texas State University, Denton, Texas. All of the subjects resided in the greater Dallas-Fort Worth area.
The sexes were equally represented, with 60 (52.2%) females and 55 (47.8%) males. The ethnic make-up of the sample was as follows: Caucasians, 90.4%; Blacks, 3.5%; Mexican-Americans, 1.7%; Orientals, 2.6%; Others/prefer not to say, 1.7%. The age range of the subjects was rather broad, 16 - 52 years ($\bar{x} = 27.9$, $SD = 7.0$).

Because of interest in comparing various occupational groups to ascertain which one was the best judge, samples were taken from numerous different groups. A group was operationally defined as a minimum of 10 individuals who were being trained in a certain profession, or who currently were employed in that profession. The following groups were established for this study, and the numbers of each are included:

- 10 psychology graduate students (clinical/counseling only)
- 10 counseling education graduate students
- 10 "experienced clinicians" (all had PhD's)
- 10 scientists (biology, chemistry, and physics)
- 10 fine arts students (all were in music)
- 10 education students
- 20 business students
- 10 secretaries
- 10 psychiatric patients
- 15 others

No complete social class information was available; however, pertinent data collected suggested middle class status for the majority. The average educational level of
the subjects' fathers was 14.0 years; mothers, 13.2 years. The subjects themselves were well-educated, with a mean educational level of 16.3 years (range = 9 - 21 years). Only one subject had less than 12 years of education and that was a ninth-grade high school student. The subjects were intellectually bright. The average Culture Fair IQ for this group was 115.2 (range = 52 - 161). This indicates that the mean IQ is at the 73 percentile. (The Culture Fair IQ test, like most IQ tests, has a mean of 100; however, its standard deviation is 24, which is larger than most standard deviations for IQ tests.)

All the psychology students were working on a PhD in either clinical or counseling psychology. All but three of them were interns at the North Texas State University (NTSU) Counseling and Testing Center. The majority (eight out of ten) were males.

The counseling education graduate students were all working on their doctorates at NTSU. An equal representation of the sexes comprised this group. Only the experienced clinicians ($\bar{x} = 37.1$ years) were older than the counseling education graduate students ($\bar{x} = 34.7$ years).

The scientists were a more diverse group than those previously mentioned. This group was represented by a research physicist with a PhD, a chemistry graduate student, and eight biology graduate students. The sex ratio consisted of six males and four females.
Individuals in the music field represented the fine arts group for this study. The subjects' interests were the following: opera, six; composition, three; instrumental music, one. All except two of the music students were graduate students. The sex ratio of this group was relatively balanced, with four males and six females.

The majority (six) of the ten education students were graduate students, although only one of them was pursuing a doctorate; however, it should be noted that the other four subjects of this group had at least a bachelor's degree in education. The areas of focus for the six graduate students and their numbers were as follows: administration, three; early childhood, one; health, one; and higher education, one. The composition of this group was three men and seven women.

The largest group for this study was the business students, which was comprised of six graduate students and fourteen undergraduates. The graduate students consisted of five doctoral students and one master's level student. Those five doctoral students' areas were as follows: management, four; research, one. The one master's level student's area of interest was personnel and industrial relations. The fourteen undergraduates were all students in a junior-level business management class. This was a male dominated group, with seventeen males and three females.

The secretarial group was entirely made up of women. This group consisted of six full-time secretaries, although
it should be noted that all ten were taking courses at the time of testing. At least five of the ten secretaries were executive level secretaries.

The group of Others consisted of fifteen individuals in widely diverse occupations. Most (eleven) of those subjects came from a Human Relations class at Richland Junior College, Dallas, Texas. The remaining four subjects consisted of two high school students and two graduate students who did not fit into the preexisting categories. The junior college students had jobs ranging from police cadet to Playboy Bunny! As might be expected, this group had the lowest mean age (22.7 years) of all ten groups. This group was composed of four males and eleven females.

The experienced clinicians' group consisted of ten individuals who had doctorates either in clinical/counseling psychology or in counseling. At the time of the study all ten were actively engaged in psychotherapy. Half of those subjects were on the staff of the NTSU Counseling and Testing Center. The other half were engaged in private practice in varying amounts of time. The clinicians averaged 8.55 years of experience in the helping professions. As might be expected, this was the oldest group in the study, with a mean age of 37.1 years. This group was comprised of six males and four females.
To add greater diversity to the subject pool, a psychiatric population was found. All ten subjects were residents of a half-way house in the Dallas-Fort Worth area. All had previously been hospitalized for a psychiatric disorder. Half of those subjects had had more than one psychiatric hospitalization. These ten subjects received the following diagnoses, and the numbers of each are reported: schizophrenia, five; manic-depressive psychoses, one; depressive neurosis, three; drug dependency, one. The mean amount of time since their release from the hospital was 3.9 months (range: one week to six months). The majority (eight) were on medication at the time they were evaluated. This was a relatively young group, averaging 24.9 years. This group was made up of two males and eight females.

**Instruments**

No known "test of experience" exists, so the following items were constructed to obtain a measure of "experience." Their respective numbers from the Personal Data Sheet, which is included in Appendix D, are included, and the rationale for their inclusion follows in parentheses.

3. Age?
   (The older one is, the more experiences in life one is likely to have.)
4. The number of children in your family while growing up?
(The assumption was made that the more siblings in a family would allow one more opportunities to learn about others.)

9. Age of your first job which required you to work at least 20 hours every week?
(Earlier employment would allow one to have more experiences in the work world.)

10. How many different jobs (both full-time and part-time) have you had in your life time?
(With more different jobs, an individual would come in contact with a greater variety of people.)

11. For how many years have you been employed in your life time?
(Those who have worked more would have had greater opportunities.)

12. How many different cities have you resided in during your life time?
(The supposition was made that for those who have lived in many different cities, the greater the opportunities to experience different cultures.)

13. How many different states/countries have you resided in during your life time?
(Similar to the above view, those who have lived in diverse cultural settings would have experienced a much greater variety of people and experiences.)
14. Compared to others my age, I have traveled:

<table>
<thead>
<tr>
<th></th>
<th>Very Much</th>
<th>An Average</th>
<th>Very Much</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less Than Most</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Less Than Most</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Average</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>More Than Most</td>
<td>10</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

(The assumption was made that more travel experiences would have broadened an individual's horizons.)

15. Did you grow up primarily in a rural or urban area? (The conjecture was made that those living in an urban area would encounter more contrasting experiences than those raised in a rural area.)

18. Highest grade level completed by you: (Those who have more educational experiences would have a broader perspective of views and ideas.)

The question could be raised concerning the associations among the experience variables. In order to investigate these, a correlation matrix was developed to show the statistical relationship among the experience variables. The reader is referred to Table 1.

If all the experience variables were measuring but one construct, high correlations would be expected among all the various experience variables. Inspection of Table 1 does not reveal that trend. The question is then raised concerning the underlying dimensions of the 10 experience measures. A principal-axis factor analysis was performed on these 10 variables. A varimax method of rotation was employed. Factoring was
<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>Number of Siblings</th>
<th>Age of First Job</th>
<th>Number of Jobs</th>
<th>Years Employed</th>
<th>Number of Cities</th>
<th>Number of States</th>
<th>Travel Exp.</th>
<th>Rural Urban</th>
<th>Educational Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.00</td>
<td>-0.18</td>
<td>-0.11</td>
<td>0.41</td>
<td>0.77</td>
<td>0.33</td>
<td>0.21</td>
<td>0.16</td>
<td>-0.04</td>
<td>0.54</td>
</tr>
<tr>
<td>Number of Siblings</td>
<td>-0.18</td>
<td>1.00</td>
<td>-0.04</td>
<td>-0.16</td>
<td>-0.11</td>
<td>-0.09</td>
<td>-0.06</td>
<td>0.02</td>
<td>-0.01</td>
<td>-0.26</td>
</tr>
<tr>
<td>Age of First Job</td>
<td>-0.11</td>
<td>-0.04</td>
<td>1.00</td>
<td>-0.25</td>
<td>-0.35</td>
<td>-0.13</td>
<td>-0.22</td>
<td>-0.04</td>
<td>-0.05</td>
<td>-0.07</td>
</tr>
<tr>
<td>Number of Jobs</td>
<td>0.41</td>
<td>-0.16</td>
<td>-0.25</td>
<td>1.00</td>
<td>0.51</td>
<td>0.30</td>
<td>0.38</td>
<td>0.25</td>
<td>-0.05</td>
<td>0.41</td>
</tr>
<tr>
<td>Years Employed</td>
<td>0.77</td>
<td>-0.11</td>
<td>-0.35</td>
<td>0.51</td>
<td>1.00</td>
<td>0.25</td>
<td>0.23</td>
<td>0.15</td>
<td>-0.07</td>
<td>0.47</td>
</tr>
<tr>
<td>Number of Cities</td>
<td>0.33</td>
<td>-0.09</td>
<td>-0.13</td>
<td>0.30</td>
<td>0.25</td>
<td>1.00</td>
<td>0.71</td>
<td>0.44</td>
<td>0.04</td>
<td>0.13</td>
</tr>
<tr>
<td>Number of States</td>
<td>0.33</td>
<td>-0.06</td>
<td>-0.22</td>
<td>0.38</td>
<td>0.23</td>
<td>0.71</td>
<td>1.00</td>
<td>0.43</td>
<td>0.06</td>
<td>0.10</td>
</tr>
<tr>
<td>Travel Experience</td>
<td>0.16</td>
<td>0.02</td>
<td>-0.04</td>
<td>0.25</td>
<td>0.15</td>
<td>0.44</td>
<td>0.43</td>
<td>1.00</td>
<td>0.06</td>
<td>0.22</td>
</tr>
<tr>
<td>Rural-urban Background</td>
<td>-0.04</td>
<td>-0.01</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.07</td>
<td>0.04</td>
<td>0.06</td>
<td>0.06</td>
<td>1.00</td>
<td>0.09</td>
</tr>
<tr>
<td>Educational Level</td>
<td>0.54</td>
<td>-0.26</td>
<td>-0.07</td>
<td>0.41</td>
<td>0.47</td>
<td>0.13</td>
<td>0.10</td>
<td>0.22</td>
<td>0.09</td>
<td>1.00</td>
</tr>
</tbody>
</table>
stopped when the eigenvalues went below 1.00. The reader can see the results of the factor analysis in Table 2.

Table 2
Rotated Factor Matrix

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>h²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.84</td>
<td>.14</td>
<td>-.07</td>
<td>-.08</td>
<td>.73</td>
</tr>
<tr>
<td>Number of Siblings</td>
<td>-.33</td>
<td>-.02</td>
<td>.61</td>
<td>-.12</td>
<td>.49</td>
</tr>
<tr>
<td>Age of First Job</td>
<td>-.33</td>
<td>-.07</td>
<td>-.77</td>
<td>-.15</td>
<td>.73</td>
</tr>
<tr>
<td>Number of Jobs</td>
<td>.64</td>
<td>.34</td>
<td>.11</td>
<td>-.06</td>
<td>.53</td>
</tr>
<tr>
<td>Years Employed</td>
<td>.86</td>
<td>.10</td>
<td>.22</td>
<td>-.09</td>
<td>.82</td>
</tr>
<tr>
<td>Number of Cities</td>
<td>.18</td>
<td>.86</td>
<td>.01</td>
<td>-.02</td>
<td>.76</td>
</tr>
<tr>
<td>Number of States/Countries</td>
<td>.14</td>
<td>.86</td>
<td>.14</td>
<td>.02</td>
<td>.78</td>
</tr>
<tr>
<td>Travel Experience</td>
<td>.09</td>
<td>.72</td>
<td>-.09</td>
<td>.07</td>
<td>.53</td>
</tr>
<tr>
<td>Rural-urban Background</td>
<td>-.05</td>
<td>.05</td>
<td>.02</td>
<td>.97</td>
<td>.94</td>
</tr>
<tr>
<td>Educational Level</td>
<td>.74</td>
<td>.05</td>
<td>-.28</td>
<td>.21</td>
<td>.68</td>
</tr>
</tbody>
</table>

Factor Variances

<table>
<thead>
<tr>
<th>Factor</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32.8</td>
<td>32.8</td>
</tr>
<tr>
<td>2</td>
<td>15.8</td>
<td>48.7</td>
</tr>
<tr>
<td>3</td>
<td>11.2</td>
<td>59.9</td>
</tr>
<tr>
<td>4</td>
<td>10.3</td>
<td>70.2</td>
</tr>
</tbody>
</table>
Inspection of Table 2 reveals that the experience measures were made up of four main factors. Factor 1, which was the biggest factor appears to represent "age." It consisted mainly of age, number of jobs, number of years employed, and educational level. The second factor was quite obviously reflective of travel experience. It consisted predominantly of number of cities resided in during life time, number of states/countries resided in during life time, and a self-evaluation of travel experience. The third factor primarily loaded on two dimensions. Those two were number of siblings in the family of origin, and age of first job. The latter loaded negatively on Factor 3. The meaning of this dimension seemed more elusive than the previous two factors. The guess was made that it represented social class. The only variable contributing to the fourth factor to any extent was rural vs. urban background.

The Culture Fair Intelligence Test (CFIQ) was employed as the measure of intelligence. Cattell (1973) has argued that this test measures fluid intelligence, rather than crystallized intelligence, as do most traditional intelligence tests. In brief, Cattell believes that fluid intelligence is strongly related to an individual's ability to perceive relationships in any material, new or old, whereas crystallized intelligence is very much loaded with general school achievement. Also, Cattell believes that fluid intelligence is largely the result of genetics, whereas crystallized intelligence is predominantly the result of learning.
The CFIQ test is a nonverbal IQ test that is comprised of four subtests. Each of the subtests requires the subject to perceive the relationships among the various shapes and figures. Like most IQ tests, the CFIQ has a mean of 100, but it has a standard deviation of 24. The large standard deviation is a result of Cattell's finding that fluid intelligence is more variable than crystallized intelligence.

Three scales of the CFIQ exist. Scale 3, which is designed for bright high school students, college students, and adults of superior intelligence, was the CFIQ scale utilized in this study. Each scale has alternate forms (A and B); however, only Form A was used in the present investigation.

The one week test-retest reliability for Scale 3, Form A, is .69. Various measures of internal consistency have been calculated. Those results, including both split-half and appropriate internal consistency formulas, yield an average internal consistency of .74 (Cattell, 1973).

Concept (construct) validity is measured by correlating the CFIQ scores with a pure measure of fluid intelligence. (The latter is a factor score.) The concept validity of Scale 3, Form A, is .81. The average concurrent validity of the CFIQ is .74. The following correlations were obtained between Scale 3, Form A, and the following well-known IQ tests: WAIS, .74; Revised Beta IQ, .76; G of the GATB, .74 (Curtis, McBrady, Clark, & Kyle, 1964).
The Eysenck Personality Questionnaire (EPQ, 1975 edition) was utilized as one measure of personality. This 90-item questionnaire has four scales: Psychoticism (P), Extraversion (E), Neuroticism (N), and Lie (L). High scorers on E tend to be outgoing, impulsive, and uninhibited. Low scorers on E, introverts, tend to be quiet, introspective, and reserved. "Emotional lability" and "over-reactivity" are measured by N. High scorers on N report many worries, anxieties, and somatic complaints. High scorers show a predisposition to develop a neurotic disorder.

The P dimension measures a predisposition for the development of a psychotic disorder (Eysenck & Eysenck, 1968). Because less is known about the meaning of P, a more thorough description is given of that construct. Eysenck and Eysenck (1975) described the high scorer on P:

A high scorer [on P], then, may be described as being solitary, not caring for people; he is often troublesome, not fitting in anywhere. He may be cruel and inhuman, lacking in feeling and empathy, and altogether insensitive. He is hostile to others, even his own kith and kin, and aggressive, even to loved ones. He has a liking for odd and unusual things, and a disregard for danger; he likes to make fools of other people, and to upset them. (p. 5-6)
L was designed to detect those who attempt to "fake good." High scorers on L tend to be socially naive and overvalue their moral worth. Less is known about this scale than the other scales developed by Eysenck and Eysenck.

Eysenck and Eysenck (1976) have presented much reliability information about their scales. None of the four scales had below a .78 reliability on a one month test-retest. Internal consistency checks were also made by them. In the data presented by Eysenck and Eysenck, none of the reliabilities for E or N were below .80; however, the internal reliability of P varied between .68 and .77.

Eysenck and Eysenck presented numerous studies (McPherson, Presley, Armstrong, & Curtis, 1974; Slade, 1975; Verma & Eysenck, 1973) to support the relationships of their P and N scales to their respective clinical syndromes. Much concurrent validity exists between the EPQ and other major personality inventories, such as the MMPI and 16PF (Cattell & Kline, 1977; Hundleby & Conner, 1968).

The Flexibility (Fx) scale of the California Psychological Inventory (CPI) was employed as the measure of cognitive complexity. As noted earlier, cognitively complex individuals are characteristically cognitively flexible. Fx was designed by Gough (1968a, 1969) to identify individuals who are flexible in their thinking, behavior, and temperament. High scorers on Fx are seen as flexible and tolerant; low scorers, rigid and authoritarian, i.e., not cognitively flexible.
Fx consists of 22 items (1 is keyed true and the other 21 false). Megargee (1972) cited much evidence concerning the reliability of this scale. For 179 freshmen women, Hase and Goldberg (1967) found a short-term reliability (1 - 4 weeks) of .49. A corrected split-half reliability of .62 was obtained on 500 males and females. A one year test-retest reliability on 234 men and women on this scale was found to be .63. A measure of internal consistency (KR-21) was calculated for high school males and females. The resulting KR-21's were .56 and .51, respectively (Megargee, 1972).

Validity information was also provided by Gough (1951). He reported significant correlations between Fx and measures of rigidity. For example, he found a correlation between the California F test (a measure of authoritarianism) and Fx of -.58 and -.36 with undergraduates and medical students, respectively. Other validity studies on Fx have been more equivocal (Megargee, 1972). Megargee noted that while Fx does correlate negatively with measures of rigidity, positive correlations with measures of flexibility have been more sparse.

The Psychological Mindedness scale (Py) from the CPI was employed to measure psychological mindedness. Gough (1968a) designed the scale to identify individuals who are psychologically oriented and insightful concerning others. He further
described the scale as the degree to which the individual is interested in, and responsible to, the inner needs, motives, and experiences of others (Gough, 1969).

Megargee (1972) cited several different measures of reliability for Py. For 179 freshman women, Hase and Goldberg (1967) found a short term reliability (1 - 4 weeks) of .74; for 200 prisoners, .53 (Gough, 1969). A one year test-retest reliability on 234 men and women was found to be .48. A corrected split-half reliability of .62 was obtained on 500 men and women. A measure of internal consistency (KR-20) was also computed on the 179 freshman women. The results indicated an internal consistency of .44.

Gough (1969) presented two different kinds of data to support the validity of Py. He found that Py correlated significantly with the Psychologist's scale of the Strong Vocational Interest Blank (ranging from .40 to .44). His second support was that psychology graduate students and those in allied fields tended to score higher than those in less psychologically-oriented occupations. Megargee (1972) has questioned the psychological mindedness interpretation of the scale, suggesting that it is more a measure of "characteristics associated with success in scientific psychology, than psychological mindedness" (p. 88).

Recent research on Py by Burkhart, Gynther, and Christian (1978) strongly suggested that this scale measures more than just scientific psychological interests. These researchers
found that high scorers on Py, much more than low scorers, were able to "more sharply discriminate the social meaning of an interpersonal communication" (p. 79). The authors concluded that their research offered a degree of construct validity to Py not previously demonstrated.

To measure the process of empathy, an abbreviated form of Hogan's Empathy Scale was employed. The original 64-item scale was constructed by Hogan (1969) to identify individuals who have good ability to empathize with others. Those 64 items came from three sources: CPI (31 items), MMPI (25 items), and the Institute of Personality Assessment and Research (IPAR) (8 items).

In the present study, an abbreviated version (39 items) of Hogan's test was employed. Those items consisted of the 31 items from the CPI and 8 items from IPAR. Hogan has cited data showing that this shortened version of his scale yielded a .92 correlation with the full scale version of his test.

Hogan has also obtained reliability information for his scale. He found a two month test-retest reliability of .84. A measure of internal consistency (KR-21) on 100 military officers was also calculated, which yielded a coefficient of .71.
Numerous validity studies have been conducted by Hogan. This scale correlated .62 with Q-sort ratings of 211 individuals' descriptions of an empathic person. Social acuity scores were made on five different groups by IPAR. Ratings made by the assessment staff at IPAR showed an average correlation between social acuity ratings and Hogan's Empathy Scale scores of .58.

Both Py and Fx are from the CPI. Hogan's Empathy Scale, in its abbreviated form, is largely (31 out of 39 items) made up of items from the CPI. Items from these three scales were all placed on one test form, rather than have subjects take three different tests. In addition to being a more efficient procedure, the assumption was made that by having items mixed together, the purposes of each of the scales would be better served, i.e., the subjects would have greater difficulty ascertaining what was being assessed. The result was a 79-item test labeled Personality Test. The subjects were instructed to answer each item either true or false. Answers were placed on an IBM answer sheet. A copy of this test is found in Appendix F.

Information was obtained concerning the internal reliabilities of each of these three scales, as was applied to the 115 subjects in this study. Cronbach's alpha was the measure of internal consistency. For Py, Cronbach's alpha was found to be .55, which was slightly higher than the figure cited by Megargee (.44). Cronbach's alpha for Fx was .80, which
was much higher than the KR-21's of .56 and .51 cited in Megargee. The measure of internal consistency for Hogan's Empathy Scale was .65, which was quite similar to that obtained by Hogan (.71).

The ability to predict stereotype accuracy is one of two identified types of person perception research. No known standardized test of stereotype accuracy has been developed; therefore, it was necessary to construct such a measure. Items selected from the MMPI, which are answered either true or false, were employed to develop such a test. These items were taken from An MMPI Handbook (Vol. 1) (Dahlstrom, Welsh, & Dahlstrom, 1972). Item endorsements, the percentage of true's and false's, were obtained from Swenson, Pearson, and Osborne (1973), who had information from 50,000 respondents at the Mayo Clinic. Sex and age norms by decade, e.g., 20-29, 30-39, 40-49, 50-59, 60-69, 70+ were reported.

Typically, the two age groups that tended to differ most from the general population average were the groups 16 - 19-years-old and the group aged 70 and over. Because these two groups were the most deviant from the average respondent, they were chosen as the two age-target groups.

Items were chosen on the basis of a rational-empirical approach. The manner in which the particular items were chosen was the following: For an item to be used, it had to
be scored in the opposite direction from the way that they majority of subjects answered it. Not only did it have to be scored in the opposite direction, but it had to differ by at least 10% from the general mean of that item. One brief example will possibly best capture the flavor of the idea: If 60% of the general Mayo Clinic population answered an item true, but only 40% of the 16 - 19-year old males did, the item was selected. The rationale behind this method of selecting items was to prevent subjects from merely projecting their own responses and thereby giving them the correct answer.

Also, the hope was that by varying both the age and sex of the target person, no group of subjects, either by age or sex, would have a projecting-the-right-answer advantage. The four groups to guess their responses were males and females, 16 - 19 years old, and males and females over 70.

By the above-mentioned method, 76 items were chosen. After selecting 76 items that were scored in the opposite direction from the modal direction, an additional 47 "obvious" items were chosen, i.e., those items were answered in the same direction by all groups. The "obvious items" and the rationally chosen items combined to make a 123-item test.

The 123 items were administered to 21 subjects. Most (15) of the subjects were either interns (graduate students in psychology or counselor education) or support personnel
(secretaries and work-study students) at the NTSU Counseling and Testing Center. Other subjects (six) were recruited from the student teaching office in the department of education at NTSU. The majority of those were support personnel (secretaries and work-study students), although two of them were doctoral students in education.

The 123-item test was scored by computer, and an item analysis was computed. Part of the function of the pretest was to reduce the number of items. It was decided that a "good item" would be defined as one with a point biserial correlation equal to or greater than + .20. Using that criterion, 29 items of the 129 met that standard. An additional analysis was done on these 29 items to ascertain the internal consistency of these 29 items. The resulting Cronbach's alpha for those 29 items was .89.

Once the 29-item test was developed, that test was administered to 23 education students who were registered for an undergraduate class in education. All those subjects were women. Those subjects were combined with the original subject pool, making a group of 44 subjects. The resulting index of internal consistency was .79.

Examining the 29 items more closely, it was noted that 22 were keyed true and only 7 false. Hence, 5 additional items, not having point biserials greater than + .20, but all
scored false, were added to the test in order to balance the number of true's and false's. One more internal consistency was measured. Again the internal consistency check was well above .70. The new 34-item test with 22 items keyed true and 12 false yielded a Cronbach's alpha of .77.

At the conclusion of the study an item analysis and measure of internal consistency were again computed on the 34-item test. The results indicated that the internal consistency was low (Cronbach's alpha = .30). Hence, the worst items from the test, i.e., those with the lowest point biserial correlations, were deleted. The resulting 25-item test had a Cronbach's alpha of .60; therefore, all subjects' scores on a measure of stereotype accuracy were based on the new 25-item test. A copy of that 34-item test, the item numbers from the MMPI, and the direction of scoring are found in Appendix H.

The Emotion Recognition Test was employed as the measure of sensitivity to individual differences. The subjects' task was to identify correctly the emotions from 36 photographs of facial expressions. The subjects had nine different categories of emotions to choose from. The nine different emotional categories of emotions and their descriptions by Izard (1971) are the following:

A. INTEREST-EXCITEMENT: Concentrating, attending, attracted, curious

B. ENJOYMENT-JOY: Glad, merry, delighted, joyful
C. **SURPRISE-STARTLE**: Sudden reaction to something unexpected, astonished

D. **DISTRESS-ANGUISH**: Sad, unhappy, miserable, feels like crying

E. **DISGUST-REVULSION**: Repugnance, aversion, distaste, sickened

F. **ANGER-RAGE**: Angry, hostile, furious, enraged

G. **SHAME-HUMILIATION**: Embarrassed, ashamed, guilty, shy

H. **FEAR-TERROR**: Scared, afraid, terrified, panicked

I. **CONTEMPT-SCORN**: Disdainful, sneering, derisive, haughty (p. 429)

Izard developed this test. From a group of 1000 photographs, he selected four pictures for each of the above nine categories. For a photo to be selected, two out of three judges had to decide that it belonged in a certain grouping. Then the photographs were presented to 30 new judges. A photograph was retained only if 70% of those 30 subjects agreed on which group it belonged. Different individuals were employed in each photograph. A copy of the answer sheet for the subjects can be found in Appendix G.

Izard subsequently showed these 36 photographs to members of different cultural groupings (Japanese, English, French, German, Swedish, Swiss, African, and American). Subjects from other cultures, like the American sample, consisted of college students. Examining the accuracy scores for the
various groups, Izard revealed that, excluding the Japanese group (different racial group from photographic subjects) and the African group (they were not tested in their native language), all other groups had, at minimum, 75% accuracy on each of the 36 photographs.

All answers to the Emotion Recognition Test were scored by hand. At the conclusion of the study, all those answer sheets were recoded on an IBM answer sheet. Each answer was dichotomized into a right or wrong category. An item analysis and measure of internal consistency were then computed. The resulting Cronbach's alpha was .68. To increase the Cronbach's alpha, the four items with the lowest point biserial correlations with the whole test score were deleted. The new test consisted of 32 pictures. All the subjects' tests were rescored, and a new Cronbach's alpha was computed on the revised test. The resulting measure of internal consistency was .71.

Procedure

Numerous different test sessions were necessary in order to complete the data collection. Prior to testing each subject, the purpose and the procedure of the study were explained to the subjects. A test packet was given to each subject. Not only were the subjects given information about
the study orally, but a written statement summarizing the purposes and procedures was included in the test packet. A copy is contained in Appendix C.

The first task for each subject after reading the instructions, and signing the permission slip to be a subject, was the Culture Fair IQ (CFIQ) Test. Because it was the only timed test, it was administered first.

At the completion of the CFIQ Test, the subjects were given the Emotion Recognition Test. Each subject, unless in a large group setting, was given the book containing the pictures. The subjects were told to study the descriptors for each of the nine emotions before venturing a guess for each picture. This test was not timed. In the case of the group administrations, the book was held up by the examiner, and the pictures were individually shown by him. At the conclusion of the group testing, subjects were given the opportunity to see any photograph again, or they were allowed to see the book itself.

After completing the Emotion Recognition Test, the subjects were given instructions to all the other tests in the test packet, which they could answer at their leisure. In addition to verbal instructions, all tests had the instructions written on them. No difficulty was encountered by any subject in comprehending the instructions by the time
they left the examiner. Typically, the subjects made a date to return the test packet within a short period of time after the initial test administration.

In four different groups of 17, 14, 8, and 6 subjects, respectively, all test data were completed in one sitting; otherwise, the subjects completed the data at their own pace. Approximately 50% of the subjects were tested individually. All the rest were tested in group settings; however, the group settings typically consisted of two individuals. The following were the group sizes larger than four individuals: 17, 14, 8, 6, and 6. All the other groups had no more than four individuals in them.

All subjects were tested by the author with the exception of the psychiatric patients. They were tested by a female graduate student who was working with them in a half-way house setting. This was done because of the belief the subjects would feel more comfortable toward someone whom they had worked with previously. Before the psychiatric patients were evaluated, the new researcher was thoroughly familiarized with the testing procedure.

**Collection and Treatment of Data**

One subject completed all data except the Eysenck Personality Questionnaire (EPQ). Rather than discard his data, the average score on each of the four scales was assigned to him. Complete data was available on all the other subjects.
All instruments employed in this study were scored by hand with the exception of the Empathy Test (measure of the stereotype other) and Personality Test, which consisted of measures of Flexibility, Psychological Mindedness, and Hogan's Empathy Scale. Those were scored by computer.

To compute the relationship between variables in Hypotheses 1 through 20, excluding Hypotheses two and three, a Pearson product-moment correlation was applied. Hypotheses two and three were both tested using a point biserial correlation. To ascertain any group differences on the two judgment tasks (Hypotheses 20 and 21), a one-way analysis of variance was computed. A Newman-Keuls test was used to ascertain any group differences. Hypotheses 22 through 25 compared those in the helping professions with those in the fine arts and sciences. A t-test was utilized to assess these differences. As the direction was indicated in all hypotheses, with the exception of Hypotheses 20 and 21, tests of significance were judged as significant at the .05 level using a one-tailed test.

Both Hypotheses 26 and 27 involved multiple correlation and the development of multiple linear regression equations. The stepwise multiple linear regression program from the Statistical Package for the Social Sciences, 2nd edition (Hie, Hull, Jenkins, Steinbrenner, & Bent, 1975), was employed
to calculate the multiple correlations. The Biomedical Computer Program, University of California, a similar program to the one employed, was described by Dickson (1970). He said that the basic process utilized to obtain the regression weights and the multiple correlations was the following:

This program computes a sequence of multilinear regression equations in a stepwise manner. At each step one variable is added to the regression equation. The variable added is the one which makes the greatest reduction in the error sum of squares. Equivalently it is the variable which has highest partial correlation with the dependent variable partialed on the variables which have already been added and equivalently it is the variable which if it were added would have the highest $F$ value. In addition, variables can be forced into the regular equation. Non-forced variables are automatically removed when their $F$ values become too small. (p. 234)

Three criteria were employed in determining when to stop adding variables:

1. When the individual beta weights were no longer significantly different from .00 at the .10 level, no additional variables were added.

2. When the variable to be added explained less than 1% of the dependent variable, no additional variables were added.
3. In no case was the variable-subject ratio allowed to be less than 3:1.

Results

In this section, data are examined as they relate to each hypothesis. Other statistical data that are judged to be germane, although not necessarily part of the original hypotheses, are also presented.

This study was designed to investigate the relationships between personality variables identified by Gordon Allport and two different kinds of person perception. One of the earlier stated goals of this study was to note general trends in relationships with measures of person perception. To meet this goal, an attempt was made to have as heterogeneous a group as possible. Consistent with that view, the decision was made that the best possible way of supporting or rejecting hypotheses would be on the available data from all 115 subjects.

Although all hypotheses are accepted or rejected on the basis of the responses of all 115 subjects, noting the correlations separately for the 55 male subjects and the 60 female subjects was believed important. Bronfenbrenner et al. (1958) have previously emphasized the importance of separation of data by sex. This was done with the thought in mind of distinguishing tendencies that may be more or less important for one of the sexes, but not the other.
Equally relevant was the decision to analyze the data by two large groups: those in the helping professions and those not in the helping professions. This was done for two reasons: (1) The Helpers typically are more test-wise than the other subjects, and they might be able to more easily "fake" their results or answer in a socially desirable direction. (2) Noting any different trends in the correlations for those in the helping professions, which might be different from those in the non-helping fields, was seen as important. Hence, data are further separated into Helpers and Non-helpers.

Because of the correlation of one of the judgment tasks with IQ, knowing what the correlations would be if IQ were partialed out seemed valuable. By making such a computation, one can even better evaluate the relationship between the numerous personality variables and the judgment tasks.

The first hypothesis states that there would be a significant positive correlation between the two types of person perception. This hypothesis was supported. Examination of Table 3 reveals a significant Pearson product-moment correlation ($r = .18, df = 115, p < .05$) between the two judgment tasks. One can see that this trend was stronger for the female subjects ($r = .24, df = 60, p < .05$) than the male subjects ($r = .08, df = 55, p > .05$).
Table 3

Pearson Product-Moment Correlations between Psychometric Variables and the Two Judgment Tasks

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>n</th>
<th>Emotion Recognition Test</th>
<th>Stereotype Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ</td>
<td>Males</td>
<td>55</td>
<td>.35**</td>
<td>-.12</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.39***</td>
<td>-.01</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.36***</td>
<td>-.08</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>.25</td>
<td>-.29</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>.38***</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>---</td>
<td>----</td>
</tr>
<tr>
<td>Psycho-</td>
<td>Males</td>
<td>55</td>
<td>.09</td>
<td>-.05</td>
</tr>
<tr>
<td>ticism</td>
<td>Females</td>
<td>60</td>
<td>-.24*</td>
<td>.04</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>-.11</td>
<td>-.02</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>-.06</td>
<td>-.23</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>-.18*</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>-.06</td>
<td>-.03</td>
</tr>
<tr>
<td>Extraversion</td>
<td>Males</td>
<td>55</td>
<td>.22*</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.27*</td>
<td>.16</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.25**</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>.00</td>
<td>.31**</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>.30**</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>.18*</td>
<td>.11</td>
</tr>
</tbody>
</table>
Table 3--Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>n</th>
<th>Emotion Recognition Test r</th>
<th>Stereotype Accuracy r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroticism</td>
<td>Males</td>
<td>55</td>
<td>-.02</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>-.19</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>-.09</td>
<td>.20**</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>.48**</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Non helpers only</td>
<td>85</td>
<td>.00</td>
<td>21*</td>
</tr>
<tr>
<td></td>
<td>IQ partialled out</td>
<td>115</td>
<td>-.04</td>
<td>.19*</td>
</tr>
<tr>
<td>Lie Scale</td>
<td>Males</td>
<td>55</td>
<td>-.30**</td>
<td>-.01</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.02</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>-.13</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>.05</td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td>Non helpers only</td>
<td>85</td>
<td>-.07</td>
<td>-.03</td>
</tr>
<tr>
<td></td>
<td>IQ partialled out</td>
<td>115</td>
<td>-.21**</td>
<td>.03</td>
</tr>
<tr>
<td>Psychological mindedness</td>
<td>Males</td>
<td>55</td>
<td>.06</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.39***</td>
<td>.16</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.17*</td>
<td>-.01</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>-.16</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>Non helpers only</td>
<td>85</td>
<td>.11</td>
<td>-.02</td>
</tr>
<tr>
<td></td>
<td>IQ partialled out</td>
<td>115</td>
<td>.08</td>
<td>.00</td>
</tr>
</tbody>
</table>
Table 3--Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>n</th>
<th>Emotion Recognition Test r</th>
<th>Stereotype Accuracy r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flexibility</td>
<td>Males</td>
<td>55</td>
<td>.21</td>
<td>-.04</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.13</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.16*</td>
<td>-.02</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>.05</td>
<td>-.23</td>
</tr>
<tr>
<td></td>
<td>Nonhelpers only</td>
<td>85</td>
<td>.05</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>IQ partialled out</td>
<td>115</td>
<td>.15</td>
<td>-.01</td>
</tr>
<tr>
<td>Hogan's Empathy Scale</td>
<td>Males</td>
<td>55</td>
<td>.29*</td>
<td>-.04</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.46***</td>
<td>.04</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.38***</td>
<td>-.01</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>.10</td>
<td>.28</td>
</tr>
<tr>
<td></td>
<td>Nonhelpers only</td>
<td>85</td>
<td>.33***</td>
<td>-.08</td>
</tr>
<tr>
<td></td>
<td>IQ partialled out</td>
<td>115</td>
<td>.33***</td>
<td>.00</td>
</tr>
<tr>
<td>Stereotype Accuracy Measure</td>
<td>Males</td>
<td>55</td>
<td>.08</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.24*</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.18*</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>-.04</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>Nonhelpers only</td>
<td>85</td>
<td>.27**</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>IQ partialled out</td>
<td>115</td>
<td>.22*</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Hypotheses two and three both dealt with sex differences on the two different judgment tasks. The prediction was made that sex and the respective judgment tasks would not be significantly correlated. A point biserial correlation was employed to test these hypotheses. One can see that there was a trend towards the female subjects doing better, but this did not reach statistical significance. Hypothesis two was rejected, \( r = -0.09, \text{df} = 115, p > .05. \) The reader can see in Table 4 that a significant point biserial correlation \( (r = -0.19, \text{df} = 115, p < .05) \) existed between sex and a measure of the stereotype other. This means that the female subjects did significantly better on the stereotype accuracy measure than did the males.
Table 4

Point Biserial Correlations between Dichotomous Variables and the Two Judgment Tasks

<table>
<thead>
<tr>
<th>Group</th>
<th>Emotion Recognition Test</th>
<th>Stereotype Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>-0.09</td>
<td>-0.19*</td>
</tr>
<tr>
<td>Females = 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helping Profession</td>
<td>-0.25**</td>
<td>0.00</td>
</tr>
<tr>
<td>Yes = 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Background</td>
<td>0.12</td>
<td>0.03</td>
</tr>
<tr>
<td>Rural = 0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban = 1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05 for a two-tailed test.

** p < .01 for a two-tailed test.

Hypothesis three was accepted. No significant point biserial correlation (r = -0.09, df = 115, p > .05) existed between sex and a measure of differential accuracy. Although Hypotheses two and three were both tested using a point biserial correlation, the utilization of a t-test would have been equally appropriate. Because of the mathematical relationship that exists between a point biserial correlation and a t-test, the identical results would be obtained which-
ever method was employed (Ferguson, 1971). Table 5 contains a summary of the results of Hypotheses two and three when a t-test was used.

Table 5
Means, Standard Deviations, and t-Values for the Two Judgment Tasks:
Males (M) vs. Females (F)

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>Means</th>
<th>SD</th>
<th>t-Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>F</td>
<td>M</td>
<td>F</td>
</tr>
<tr>
<td>Differential Accuracy</td>
<td>55</td>
<td>60</td>
<td>23.5</td>
<td>24.3</td>
</tr>
<tr>
<td>Stereotype Accuracy</td>
<td>55</td>
<td>60</td>
<td>18.4</td>
<td>19.6</td>
</tr>
</tbody>
</table>

* p < .05 for a two-tailed test.

Looking at Table 5, one can see that the results were identical to the point biserial results. Hypothesis two was accepted; hypothesis three, rejected. Females were significantly better judges of stereotype accuracy than males, t = 2.07, df = 113, p < .05. No sex differences were noted on the measure of differential accuracy.

The relationship between IQ and person perception was studied in Hypotheses four and five. Hypothesis four (that a significant positive correlation would exist between intelligence and stereotype accuracy) was not supported, r = -.08, df = 115, p > .05.
The fifth research hypothesis (that there would be a significant positive correlation between intelligence and differential accuracy) received strong empirical support, \( r = .36, \ df = 115, \ p < .001 \). Inspection of Table 3 indicates that intelligence was positively related to doing well on the Emotion Recognition Test no matter what group was being considered. And, in all but one case, the correlations reached statistical significance.

Hypotheses six and seven evaluated the relationship between cognitive complexity/flexibility and the two judgment tasks. Hypothesis six stated the following: "There will be a significant positive correlation between cognitive complexity and the stereotype other." This was not supported by the data, \( r = .02, \ df = 115, \ p < .05 \).

Hypothesis seven stated that a significant positive correlation would exist between cognitive complexity and the ability to accurately recognize facial expressions. Checking Table 3, one can see that this hypothesis was upheld by the data. A significant positive correlation was found between the two variables, \( r = .16, \ df = 115, \ p < .05 \).

The eighth and ninth research hypotheses dealt with psychological mindedness. Hypothesis eight stated that there would be a significant positive correlation between Psychological Mindedness and stereotype accuracy. This hypothesis was not corroborated by the data, \( r = -.01, \ df = 115, \ p > .05 \).
The ninth research hypothesis (that a significant positive correlation would exist between Psychological Mindedness and the Emotion Recognition Test) received empirical support, \( r = .17, \, df = 115, \, p < .05 \). Examination of Table 3 rather obviously indicates that this relationship was stronger for the female subjects \( (r = .39, \, df = 60, \, p < .001) \) than for the male subjects \( (r = .06, \, df = 55, \, p > .05) \).

The next four hypotheses were concerned with the effects of psychopathology. Research Hypothesis 10 (that there would be a significant negative correlation between Neuroticism and stereotype accuracy) was not sustained by the data, \( r = .20, \, df = 115, \, p < .01 \). As can be seen in Table 3, the correlation was significant, but it was not in the predicted direction. That is, those high on Neuroticism tended to be better judges of stereotype accuracy than those low on the same measure.

Hypothesis 11 stated the following: "There will be a significant negative correlation between Neuroticism and differential accuracy." Analysis given in Table 3 indicates that this hypothesis failed to receive support, \( r = -.09, \, df = 115, \, p > .05 \). One of the more intriguing findings concerning the relationship between Neuroticism and the ability to do well on the Emotion Recognition Test involves the helper group. For that group, Neuroticism was significantly correlated \( (r = .48, \, df = 30, \, p < .01) \) with doing well on the Emotion Recognition Test.
Because one of the symptoms of a psychotic disorder is a perceptual disorder, one would especially think that the psychoticism measure would be negatively related to the two judgment tasks. This led to Hypothesis 12: "There will be a significant negative correlation between Psychoticism and stereotype accuracy." This hypothesis was rejected, \( r = -.02, \) df = 115, \( p > .05. \)

Hypothesis 13 dealt with the relationship between Psychoticism and scores on the Emotion Recognition Test. The hypothesis was made that a significant correlation would exist between those two variables. That hypothesis was rejected, \( r = -.11, \) df = 115, \( p > .05. \) A careful study of Table 3 indicates, however, that a significant negative correlation (\( r = .24, \) df = 60, \( p < .05. \)) was found for the female subjects, but not the male subjects (\( r = .09, \) df = 55, \( p > .05. \)).

Hypotheses 10 through 14 dealt with the psychometric measures of psychopathology and their relationship to person perception. Additional information on this topic is gained from observing the performance of the Psychiatric Patients' group in this study. That data are discussed under the results for Hypotheses 20 and 21.

Historically, one of the most important constructs in psychology is extraversion-introversion. The relationship between extraversion and person perception was explored in
Hypotheses 14 and 15. Hypothesis 14 stated that there would be a significant negative correlation between Extraversion and stereotype accuracy. This was not supported by the data, $r = .09, df = 115, p > .05$.

The fifteenth research hypothesis (that a significant negative correlation would exist between Extraversion and differential accuracy) was not given empirical support; however, a significant positive correlation ($r = .25, df = 115, p < .001$) was found between Extraversion and scores on the Emotional Recognition Test. Table 3 reveals that this trend was true for both males ($r = .22, df = 55, p < .05$) and females ($r = .27, df = 60, p < .05$). Checking the different groups in Table 3 more closely, one can see that higher scores on Extraversion are more related to being a good judge of facial expression than lower scores on Extraversion (introversion).

Hypothesis 16 stated the following: "There will be a significant positive correlation between an empathic style and stereotype accuracy." (An empathic style was measured by Hogan's Empathy Scale.) This hypothesis was rejected, $r = -.01, df = 115, p > .05$.

The seventeenth research hypothesis (that a significant positive relationship existed between an empathic style and differential accuracy) received strong empirical support, $r = .38, df = 115, p < .001$. This correlation was the very
highest one found of the psychometric variables. Inspection of Table 3 reveals that this relationship was statistically significant for all the groups studied except one.

One of the most interesting variables to be studied was that of years of experience in the helping professions. Hypotheses 18 and 19 dealt with this issue. Hypothesis 18 stated the following: "For those in the helping professions, there will be a significant positive correlation between years of experience in the helping professions and stereotype accuracy." No support was given for this hypothesis, \( r = -0.21, df = 30, p > 0.05 \). A review of Table 3 enables the reader to better understand the relationship between person perception and years of experience in the helping profession.

Hypothesis 19 stated the following: "For those in the helping professions, there will be a significant positive correlation between years of experience in the helping professions and differential accuracy." This hypothesis was rejected, \( r = -0.15, df = 30, p > 0.05 \).

Previously, Allport's belief (1961) that a person with a wide range of experience was a better judge of others than one with a limited number of experiences was noted. No known "test of experience" exists. Hence, a number of items that were judged to have face validity to a measure of experience were constructed.
No specific hypotheses were made with regard to the nine experiential variables; however, some of the present findings seemed promising, and therefore are discussed. Table 6 summarizes the Pearson product-moment correlations between the various measures of experience and the two judgment tasks.

Table 6
Pearson Product-Moment Correlations between Measures of Experience and the Two Judgment Tasks

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>n</th>
<th>Emotion Recognition Test r</th>
<th>Stereotype Accuracy r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Males</td>
<td>55</td>
<td>.16</td>
<td>-.04</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.30**</td>
<td>-.09</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.24**</td>
<td>-.08</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>-.13</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>.18*</td>
<td>-.14</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>.26**</td>
<td>.08</td>
</tr>
<tr>
<td>Number of Siblings</td>
<td>Males</td>
<td>55</td>
<td>-.46***</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>-.13</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>-.28***</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>.21</td>
<td>.23</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>-.36***</td>
<td>-.04</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>-.26**</td>
<td>.04</td>
</tr>
<tr>
<td>Age of First Job</td>
<td>Males</td>
<td>55</td>
<td>.11</td>
<td>.22*</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.22*</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.18*</td>
<td>.26**</td>
</tr>
</tbody>
</table>
Table 6--Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>n</th>
<th>Emotion Recognition Test r</th>
<th>Stereotype Accuracy r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Different Jobs</td>
<td>Males</td>
<td>55</td>
<td>.25*</td>
<td>-.04</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.11</td>
<td>.05</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.15</td>
<td>-.02</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>-.20</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>.13</td>
<td>-.05</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>.13</td>
<td>-.02</td>
</tr>
<tr>
<td>Number of Years Employed</td>
<td>Males</td>
<td>55</td>
<td>.07</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.11</td>
<td>-.25*</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.07</td>
<td>-.12</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>-.26</td>
<td>-.01</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>.03</td>
<td>-.18*</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>.05</td>
<td>-.11</td>
</tr>
<tr>
<td>Number of Cities Resided in During Life Time</td>
<td>Males</td>
<td>55</td>
<td>.13</td>
<td>-.16</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.31**</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.22**</td>
<td>-.03</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>-.15</td>
<td>.26</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>.23**</td>
<td>-.11</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>.18*</td>
<td>-.02</td>
</tr>
</tbody>
</table>
Table 6--Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Group</th>
<th>n</th>
<th>Emotion Recognition Test r</th>
<th>Stereotype Accuracy r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of States/Countries Resided in During Life Time</td>
<td>Males</td>
<td>55</td>
<td>-.10</td>
<td>-.17</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.15</td>
<td>.02</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.01</td>
<td>-.11</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>-.15</td>
<td>-.05</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>.01</td>
<td>-.13</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>-.03</td>
<td>-.10</td>
</tr>
<tr>
<td>Travel Experience</td>
<td>Males</td>
<td>55</td>
<td>.12</td>
<td>-.12</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.41***</td>
<td>.18</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.28***</td>
<td>-.01</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>.08</td>
<td>-.28</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>.27***</td>
<td>-.09</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>.19*</td>
<td>-.04</td>
</tr>
<tr>
<td>Educational Level</td>
<td>Males</td>
<td>55</td>
<td>.51***</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>Females</td>
<td>60</td>
<td>.33**</td>
<td>-.21</td>
</tr>
<tr>
<td></td>
<td>All 115 Subjects</td>
<td>115</td>
<td>.36***</td>
<td>-.12</td>
</tr>
<tr>
<td></td>
<td>Helpers only</td>
<td>30</td>
<td>.05</td>
<td>-.09</td>
</tr>
<tr>
<td></td>
<td>Non-helpers only</td>
<td>85</td>
<td>.25**</td>
<td>-.14</td>
</tr>
<tr>
<td></td>
<td>IQ partialed out</td>
<td>115</td>
<td>.31***</td>
<td>-.11</td>
</tr>
</tbody>
</table>

* p < .05 for a one-tailed test.
** p < .01 for a one-tailed test.
*** p < .001 for a one-tailed test.
Allport (1961) believed that a good judge of others first of all needed maturity, which means not only an accumulation of years, but a rich store of experience. Consistent with Allport's view, age was significantly positively correlated \((r = .24, df = 115, p < .01)\) with the ability to judge facial expressions. The objection might justifiably be raised that the older subjects in this study were the better educated subjects. Hence, the brighter subjects would be the older subjects; however, even with IQ partialed out, the relationship between age and differential accuracy still was statistically significant \((r = .26, df = 115, p < .01)\). Age was not found to be related to being a good judge of the stereotype other.

More siblings in a family would seem to offer more opportunities to learn about others. The present results were in sharp contradiction to that. The fewer the siblings in the family, the greater the likelihood that an individual would be a good judge of facial expressions, \(r = -.28, df = 115, p < .001\). This trend was far more pronounced in males \((r = -.46, df = 55, p < .001)\) than females \((r = -.13, df = 60, p > .05)\). No similar effect was observed between number of siblings and being a good judge of stereotype accuracy.

Age of first job was thought to be inversely related to measures of person perception. That is, those going to work earlier in life would have greater range of experiences than
those who were more "protected." The data did not support this view. In fact, the relationship was exactly the opposite: Those who went to work later were better judges of facial expressions than those who went to work earlier. Inspection of Table 6 indicates that this trend is true for both sexes, but it was more pronounced for the females ($r = .22$, $df = 60$, $p < .05$) than the males ($r = .11$, $df = 55$, $p > .05$).

The question could be raised that social class might have been a factor in this result. As IQ is correlated with social class, the decision was made to partial out IQ to see if this relationship still existed. It did. The relationship between age of first job and the ability to accurately judge facial expressions with IQ partialed out increased, $r = .30$, $df = 115$, $p < .001$. Age of first job was also noted to be significantly related to the ability to be a good judge of the stereotype other, $r = .26$, $df = 115$, $p < .01$.

Three different measures of travel experience were obtained: number of different cities resided in during life time, number of different states/countries resided in during life time, and a self-evaluation of travel experience. The last measure was based on a Likert-type scale which had a continuum ranging from 1 - 11. As one can see in Table 6, the correlations were higher for the females than the males.
Two of those three correlations reached statistical significance for the females, but not the males. For females, the relationship between number of cities resided in during life time and ability to accurately judge facial expressions was correlated significantly, $r = .31, df = 60, p < .01$. Similarly, for the female subjects, the relationship between the Emotion Recognition Test and a self-evaluation of one's travel experience was highly significant, $r = .41, df = 60, p < .001$.

Educational level was obviously strongly related ($r = .35, df = 115, p < .001$) to doing well on a measure of sensitivity to individual differences. Interestingly enough, this trend was stronger for the males ($r = .51, df = 55, p < .001$) than the females ($r = .33, df = 60, p < .01$). Even when IQ was partialed out, the relationship between educational level and scores on the Emotion Recognition Test were still significantly related ($r = .31, df = 115, p < .001$).

One of the most relevant issues in this study was the comparison of the various groups on the two judgment tests. Hypothesis 20 stated that there would be a significant difference in the study among the various groups on a measure of stereotype accuracy. Table 7 lists the various groups in the study, their means, and standard deviations on the measure of stereotype accuracy. Examination of Table 7 shows that the worst judges and the best judges on a measure of stereo-
type accuracy were those in the helping professions! The reader's attention is directed to the position of the Psychiatric Patients, as they were the second best judges.

Table 7
Means and Standard Deviations for the Various Groups on the Stereotype Accuracy Measure

<table>
<thead>
<tr>
<th>Name of Group</th>
<th>n</th>
<th>Means</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Counselor Education Students</td>
<td>10</td>
<td>21.70</td>
<td>1.64</td>
</tr>
<tr>
<td>Psychiatric Patients</td>
<td>10</td>
<td>20.60</td>
<td>3.75</td>
</tr>
<tr>
<td>Scientists</td>
<td>10</td>
<td>19.90</td>
<td>2.18</td>
</tr>
<tr>
<td>Fine Arts Students</td>
<td>10</td>
<td>19.60</td>
<td>2.55</td>
</tr>
<tr>
<td>Others</td>
<td>15</td>
<td>19.27</td>
<td>2.22</td>
</tr>
<tr>
<td>Secretaries</td>
<td>10</td>
<td>19.20</td>
<td>2.66</td>
</tr>
<tr>
<td>Education Students</td>
<td>10</td>
<td>18.30</td>
<td>2.45</td>
</tr>
<tr>
<td>Business Students</td>
<td>20</td>
<td>17.95</td>
<td>3.38</td>
</tr>
<tr>
<td>Experienced Clinicians</td>
<td>10</td>
<td>17.40</td>
<td>4.09</td>
</tr>
<tr>
<td>Psychology Students</td>
<td>10</td>
<td>17.30</td>
<td>2.67</td>
</tr>
<tr>
<td>Totals</td>
<td>115</td>
<td>19.03</td>
<td>3.05</td>
</tr>
<tr>
<td>Non-helpers</td>
<td>85</td>
<td>19.11</td>
<td>2.88</td>
</tr>
<tr>
<td>Helpers</td>
<td>30</td>
<td>18.80</td>
<td>3.55</td>
</tr>
</tbody>
</table>
Hypothesis 20 was confirmed. Significant differences existed \( (F = 2.59, \text{df} = 9, 105, p < .01) \) among the groups. Despite the fact that a significant \( F \) was found, it should be noted that the results of Bartlett's Test, a measure of the homogeneity of variance, indicated that a significant difference existed among the groups' variances. This is mentioned, as it violates one of the assumptions underlying analysis of variance. When a Newman-Keuls test was employed to compare pairs of means, three significant differences were found. The Counselor Education Students were significantly \( (p < .05) \) better than the Psychology Students, the Experienced Clinicians, and the Business Students. No other comparisons indicated significant differences.

Hypothesis 21 stated the following: "The various groups in the study will differ significantly on a measure of differential accuracy." Hypothesis 21 was supported. A significant difference \( (F = 2.80, \text{df} = 9, 105, p < .01) \) was noted among the groups' means on the Emotion Recognition Test. When a comparison of pairs of means was conducted employing the Newman-Keuls test, only one significant difference was found. The Psychology Students did significantly \( (p < .05) \) better than the Business Students.

Table 8 lists the various groups of the study, their means, and standard deviations on the Emotion Recognition Test. Of considerable interest is the showing of the groups
Table 8
Means and Standard Deviations for the Various Groups on the Emotion Recognition Test

<table>
<thead>
<tr>
<th>Name of Group</th>
<th>n</th>
<th>Means</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychology Students</td>
<td>10</td>
<td>26.90</td>
<td>3.45</td>
</tr>
<tr>
<td>Experienced Clinicians</td>
<td>10</td>
<td>26.30</td>
<td>1.77</td>
</tr>
<tr>
<td>Education Students</td>
<td>10</td>
<td>25.40</td>
<td>3.60</td>
</tr>
<tr>
<td>Counselor Education Students</td>
<td>10</td>
<td>24.90</td>
<td>1.52</td>
</tr>
<tr>
<td>Secretaries</td>
<td>10</td>
<td>24.80</td>
<td>4.18</td>
</tr>
<tr>
<td>Scientists</td>
<td>10</td>
<td>24.40</td>
<td>2.84</td>
</tr>
<tr>
<td>Fine Arts Students</td>
<td>10</td>
<td>24.40</td>
<td>2.72</td>
</tr>
<tr>
<td>Others</td>
<td>15</td>
<td>22.27</td>
<td>4.96</td>
</tr>
<tr>
<td>Business Students</td>
<td>20</td>
<td>21.80</td>
<td>4.10</td>
</tr>
<tr>
<td>Psychiatric Patients</td>
<td>10</td>
<td>21.20</td>
<td>6.73</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>115</td>
<td>23.94</td>
<td>4.22</td>
</tr>
<tr>
<td>Helpers</td>
<td>30</td>
<td>26.04</td>
<td>2.47</td>
</tr>
<tr>
<td>Non-helpers</td>
<td>85</td>
<td>23.20</td>
<td>4.47</td>
</tr>
</tbody>
</table>

which are in the social occupational categories, i.e., Counselor Education Students, Psychology Students, Education Students, and the Experienced Clinicians. Those four groups, all oriented towards helping people, had the very highest
scores. Note the very poor showing of the Psychiatric Patients on this task. These results suggest that those who are mentally ill have less acuity in accurately recognizing emotions in others. Cause or effect?

Although not hypothesized, one very obvious comparison is that between the Helpers and the Non-helpers on the two judgment tasks. The means, standard deviations, and t-values for both groups on the Emotion Recognition Test and the stereotype accuracy measure are listed in Table 9. The Helpers were significantly ($t = 3.29$, df = 113, $p < .01$) better judges of facial expressions than the Non-helpers. No significant difference was noted on the stereotype accuracy measure, $t = .47$, df = 113, $p > .05$.

Table 9

<table>
<thead>
<tr>
<th>Variable</th>
<th>Means</th>
<th>SD</th>
<th>t-Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Helpers (30)</td>
<td>Non-helpers (85)</td>
<td>Helpers (30)</td>
</tr>
<tr>
<td>Emotion Recognition Test</td>
<td>26.04</td>
<td>23.20</td>
<td>2.47</td>
</tr>
<tr>
<td>Stereotype Accuracy</td>
<td>18.80</td>
<td>19.11</td>
<td>3.55</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are n's.

* $p < .01$ for a two-tailed test.
Certainly one of the most intriguing comparisons has to be between the Helpers and the Psychiatric Patients' group! In Table 10, the reader can see the means, standard deviations, and t-values for both groups on the differential and stereotype accuracy measures. No significant difference ($t = 1.04$, $df = 38$, $p > .05$) was found between the two groups on a measure of stereotype accuracy. However, the Helpers were significantly better judges of emotional expressions than the Psychiatric Patients, $t = 2.55$, $df = 38$, $p < .05$.

Table 10

Means, Standard Deviations, and t-Values between the Helpers and Psychiatric Patients on the Two Judgment Tasks

<table>
<thead>
<tr>
<th>Variable</th>
<th>Means</th>
<th>SD</th>
<th>t-Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Helpers (30)</td>
<td>Patients (10)</td>
<td>Helpers (30)</td>
</tr>
<tr>
<td>Emotion Recognition Test</td>
<td>26.04</td>
<td>21.10</td>
<td>2.47</td>
</tr>
<tr>
<td>Stereotype Accuracy</td>
<td>18.80</td>
<td>20.60</td>
<td>3.55</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are n's.

* $p < .05$ for a two-tailed test.

Hypothesis 22 stated that people in the helping professions would be significantly better judges than those in the fine arts on stereotype accuracy. This hypothesis was
rejected. Inspection of Table 11, which contains the means, standard deviations, and t-values for the two groups, reveals that no significant difference existed between the two groups, $t = .50$, $df = 38$, $p > .05$.

Research Hypothesis 23 (that people in the helping professions would be significantly better judges than those in the fine arts on a measure of differential accuracy) was not accepted, $t = 1.34$, $df = 38$, $p > .05$. One can see the means, standard deviations, and t-values for the two groups in Table 11.

Table 11

<table>
<thead>
<tr>
<th>Variable</th>
<th>Means</th>
<th>SD</th>
<th>t-Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Helpers</td>
<td>Fine Arts</td>
<td>Helpers</td>
</tr>
<tr>
<td>Emotion Recognition Test</td>
<td>26.04</td>
<td>24.40</td>
<td>2.47</td>
</tr>
<tr>
<td>Stereotype Accuracy</td>
<td>18.80</td>
<td>19.60</td>
<td>3.55</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are n's.

Hypothesis 24 compared the Helpers with the Scientists on a measure of stereotype accuracy. The prediction was that the former would score significantly higher than the latter. Checking Table 12, one can see the means, standard
deviations, and t-values for both groups. Hypothesis 24 was not upheld, $t = .69$, $df = 38$, $p > .05$.

Hypothesis 25, which stated that Helpers would be significantly better judges than those in the sciences on a measure of differential accuracy, was not given empirical support, $t = 1.33$, $df = 38$, $p > .05$. The means, standard deviations, and t-values for both groups are found in Table 12.

Table 12

Means, Standard Deviations, and t-Values between the Helpers and Scientists' Group on the Two Judgment Tasks

<table>
<thead>
<tr>
<th>Variable</th>
<th>Means</th>
<th>SD</th>
<th>t-Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Helpers</td>
<td>Scientists</td>
<td>Helpers</td>
</tr>
<tr>
<td>Emotion Recognition Test</td>
<td>26.04</td>
<td>24.40</td>
<td>2.47</td>
</tr>
<tr>
<td>Stereotype Accuracy Measure</td>
<td>18.80</td>
<td>19.90</td>
<td>3.55</td>
</tr>
</tbody>
</table>

Note: Numbers in parentheses are n's.

Hypothesis 26 stated that there would be a significant multiple correlation ($R$) between the predictor variables and the criterion variable of scores on a measure of stereotype accuracy. The psychometric predictor variables consisted of measures on the following: IQ, psychoticism, extraversion, neuroticism, lie, psychological mindedness, flexibility, and
empathic style. The experience predictor variables consisted of the following: age, number of states/countries resided in during lifetime, number of cities resided in during lifetime, travel experience, number of different jobs, number of years employed, age of first job, and rural-urban background. A review of Table 13 shows these results:

Table 13
Summary Statistics for Multiple Correlation Criterion: Stereotype Accuracy Measure for All Subjects

<table>
<thead>
<tr>
<th>ANOVA for the Model</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>df</td>
<td>Sum of Squares</td>
<td>Mean Squares</td>
<td>F</td>
</tr>
<tr>
<td>Regression</td>
<td>2</td>
<td>99.31</td>
<td>49.65</td>
<td>5.77***</td>
</tr>
<tr>
<td>Residual</td>
<td>112</td>
<td>963.61</td>
<td>8.60</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANOVA for the Independent Variables</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent Variables</td>
<td>Raw Score</td>
<td>z-score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variables</td>
<td>Beta Weights</td>
<td>Beta Weights</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td>Age of first job</td>
<td>.34</td>
<td>.23</td>
<td>6.40**</td>
<td></td>
</tr>
<tr>
<td>Neuroticism</td>
<td>.08</td>
<td>.16</td>
<td>3.16*</td>
<td></td>
</tr>
<tr>
<td>(Constant)</td>
<td>12.47</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 13--Continued

Summary Table for Multiple Regression

\[
\begin{align*}
R &= .305 \\
R \text{ Square} &= .093 \\
\text{Standard Error} &= 2.933
\end{align*}
\]

Hypothesis 26 was sustained, \( R = .31 \). This \( R \) was statistically significant, \( F = 5.77, df = 2, 112, p < .01 \). With an \( R \) of .31, there is an \( R \) square of .09, which means that 9% of the variance has been accounted for by the two predictors. Table 13 includes the summary statistics for \( R \), the regression weights for both the raw score beta weights, and the \( z \)-score beta weights.

Hypothesis 26 was designed to be tested by using data from all 115 subjects; however, the decision was made that additional information could be obtained by analyzing the data separately by sex. For the female subjects no statistically significant \( R \) was found between the predictor variables and the criterion score. Similarly, no statistically significant \( R \) was found for the male subjects.

The results of Hypothesis 27 can be seen in Table 14. This hypothesis stated that there would be a significant \( R \) between a combination of both psychometric predictor variables (IQ, psychoticism, extraversion, neuroticism, lie scale, psychological mindedness, flexibility, empathic style); experience variables (age, number of states/countries resided
Table 14
Summary Statistics for Multiple Correlation
Criterion: Emotion Recognition Test
for All Subjects

ANOVA for the Model

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>6</td>
<td>897.09</td>
<td>149.52</td>
<td>14.20***</td>
</tr>
<tr>
<td>Residual</td>
<td>108</td>
<td>1137.48</td>
<td>10.53</td>
<td></td>
</tr>
</tbody>
</table>

*** \( p < .001 \)

ANOVA for the Independent Variables

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Raw Score Beta Weights</th>
<th>z-Score Beta Weights</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hogan's Empathy Scale</td>
<td>.33</td>
<td>.37</td>
<td>22.37***</td>
</tr>
<tr>
<td>IQ</td>
<td>.06</td>
<td>.33</td>
<td>19.12***</td>
</tr>
<tr>
<td>Age of First Job</td>
<td>.67</td>
<td>.33</td>
<td>17.80***</td>
</tr>
<tr>
<td>Number of Siblings</td>
<td>-.55</td>
<td>-.22</td>
<td>8.92**</td>
</tr>
<tr>
<td>Number of Towns Resided in</td>
<td>.34</td>
<td>.28</td>
<td>7.32**</td>
</tr>
<tr>
<td>Number of States Resided in</td>
<td>-.54</td>
<td>-.25</td>
<td>6.04*</td>
</tr>
<tr>
<td>(Constant)</td>
<td>-.80</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** \( p < .001 \)
** \( p < .01 \)
* \( p < .05 \)
Table 14--Continued

Summary Table for Multiple Regression

<table>
<thead>
<tr>
<th>R</th>
<th>R Square</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>.664</td>
<td>.441</td>
<td>3.245</td>
</tr>
</tbody>
</table>

in during life time, number of cities resided in during life time, travel experience, number of different jobs, number of years employed, age of first job, rural-urban background); and the criterion variable of scores on the Emotion Recognition Test.

This hypothesis was given strong empirical support, R = .66. As one can see in Table 14, R was statistically significant, F = 14.20, df = 6, 108, p < .001. An R of .66 yields an R square of .44, which means that 44% of the variance of the Emotion Recognition Test was accounted for by the six predictors. Table 14 includes the summary statistics for R, the regression weights for both the raw score multiple linear regression equation, and the z-score multiple linear regression equation.

Although Hypothesis 27 was made for all 115 subjects, examination of the data separately by sex was done in hopes of obtaining greater clarification. Hence, the data were reanalyzed using the identical predictors and criterion as before. The results for the female subjects are examined first, and then the data for the male subjects are presented.
For the female subjects alone (n = 60), Hypothesis 27 was corroborated, \( R = .77 \). Table 15 indicates that \( R \) was statistically significant, \( F = 11.08, \text{df} = 7, 52, p < .001 \). With an \( R \) of .77, this yields an \( R \) square of .60, which means that 60% of the variance of the Emotion Recognition Test was accounted for by the seven predictor variables. Table 15 includes the summary statistics for \( R \), the regression weights for the raw score multiple linear regression equation, and the \( z \)-score multiple linear regression equation.

Table 15
Summary Statistics for Multiple Correlation
Criterion: Emotion Recognition Test
for Female Subjects Only

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>7</td>
<td>736.76</td>
<td>105.25</td>
<td>11.08***</td>
</tr>
<tr>
<td>Residual</td>
<td>52</td>
<td>493.84</td>
<td>9.50</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Raw Score Beta Weights</th>
<th>z-score Beta Weights</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hogan's Empathy Scale</td>
<td>.36</td>
<td>.41</td>
<td>15.14***</td>
</tr>
<tr>
<td>IQ</td>
<td>.06</td>
<td>.34</td>
<td>11.30**</td>
</tr>
<tr>
<td>Age of First Job</td>
<td>1.14</td>
<td>.43</td>
<td>21.56***</td>
</tr>
<tr>
<td>Number of Towns Resided in</td>
<td>.43</td>
<td>.33</td>
<td>12.39***</td>
</tr>
<tr>
<td>Psychological Mindedness</td>
<td>.45</td>
<td>.30</td>
<td>7.43**</td>
</tr>
</tbody>
</table>
### Table 15--Continued

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Raw Score Beta Weights</th>
<th>z-Score Beta Weights</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroticism</td>
<td>.17</td>
<td>.21</td>
<td>3.50*</td>
</tr>
<tr>
<td>Lie Scale</td>
<td>.34</td>
<td>.19</td>
<td>3.34*</td>
</tr>
<tr>
<td>(Constant)</td>
<td>-22.30</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* * p < .10  
** p < .01  
*** p < .001

Summary Table for Multiple Regression

R = .773  
R Square = .598  
Standard Error = 3.081

Hypothesis 27 was given support for the male subjects (n = 55), R = .62. Viewing Table 16, one can see that R was statistically significant, F = 16.48, df = 2, 52, p < .001. With an R of .62, this gives an R square of .39, which means that 39% of the variance of the Emotion Recognition Test was accounted for by the two predictor variables. Table 16 gives the summary statistics for R, the regression weights for
the raw score multiple linear regression equation, and the
z-score multiple linear regression equation.

Table 16
Summary Statistics for Multiple Correlation
Criterion: Emotion Recognition Test
for All Male Subjects

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>Sum of Squares</th>
<th>Mean Squares</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>2</td>
<td>305.55</td>
<td>152.77</td>
<td>16.48***</td>
</tr>
<tr>
<td>Residual</td>
<td>52</td>
<td>482.09</td>
<td>9.27</td>
<td></td>
</tr>
</tbody>
</table>

ANOVA for the Independent Variables

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Raw Score Beta Weights</th>
<th>z-Score Beta Weights</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Educational Level</td>
<td>.59</td>
<td>.43</td>
<td>14.80***</td>
</tr>
<tr>
<td>Number of Siblings</td>
<td>-.81</td>
<td>-.37</td>
<td>10.38***</td>
</tr>
<tr>
<td>(Constant)</td>
<td></td>
<td>16.16</td>
<td></td>
</tr>
</tbody>
</table>

** p < .01
*** p < .001

Summary Table for Multiple Regression

R = .622
R Square = .387
Standard Error = .304

One other attempt to view some of the important variables
was to divide the subjects into two artificial dichotomies of
good judges vs. bad judges. A good judge on the Emotion Recognition Test was arbitrarily defined as one who scored 28 or greater on that test; a poor judge, one who scored 16 or lower on the same test. The top 19 scorers (top 17%) on the Emotion Recognition Test were designated as the best judges, and the worst judges were those who scored the lowest 18 (bottom 16%), as shown in Table 17.

Table 17

Means, Standard Deviations, and t-Values between the 19 Besta Judges and the 18b Worst Judges on the Emotion Recognition Test (ERT)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Good (19)</th>
<th>Poor (18)</th>
<th>SD Good (19)</th>
<th>SD Poor (18)</th>
<th>t-Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERT</td>
<td>28.74</td>
<td>16.06</td>
<td>1.05</td>
<td>3.23</td>
<td>16.27***</td>
</tr>
<tr>
<td>IQ</td>
<td>123.05</td>
<td>100.83</td>
<td>20.62</td>
<td>31.04</td>
<td>2.58*</td>
</tr>
<tr>
<td>Psychoticism</td>
<td>2.63</td>
<td>3.55</td>
<td>2.27</td>
<td>4.30</td>
<td>.82</td>
</tr>
<tr>
<td>Extraversion</td>
<td>14.84</td>
<td>12.44</td>
<td>4.07</td>
<td>5.64</td>
<td>1.49</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>10.10</td>
<td>11.44</td>
<td>5.29</td>
<td>6.44</td>
<td>.69</td>
</tr>
<tr>
<td>Lie Scale</td>
<td>4.95</td>
<td>6.44</td>
<td>1.61</td>
<td>4.25</td>
<td>1.43</td>
</tr>
<tr>
<td>Psychological Mindedness</td>
<td>12.63</td>
<td>11.61</td>
<td>3.09</td>
<td>7.78</td>
<td>.53</td>
</tr>
<tr>
<td>Flexibility</td>
<td>13.42</td>
<td>10.61</td>
<td>3.96</td>
<td>4.03</td>
<td>2.14*</td>
</tr>
<tr>
<td>Hogan's Empathy Scale</td>
<td>27.84</td>
<td>22.50</td>
<td>4.01</td>
<td>5.68</td>
<td>3.32**</td>
</tr>
<tr>
<td>Sex</td>
<td>.32</td>
<td>.50</td>
<td>.48</td>
<td>.51</td>
<td>1.13</td>
</tr>
<tr>
<td>Age</td>
<td>29.68</td>
<td>25.06</td>
<td>6.70</td>
<td>6.96</td>
<td>2.06*</td>
</tr>
</tbody>
</table>
Table 17--Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Means</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td></td>
<td>(19)</td>
<td>(18)</td>
</tr>
<tr>
<td>Number of siblings</td>
<td>3.11</td>
<td>4.22</td>
</tr>
<tr>
<td>Age of first job</td>
<td>16.89</td>
<td>15.72</td>
</tr>
<tr>
<td>Number of different jobs</td>
<td>9.58</td>
<td>7.22</td>
</tr>
<tr>
<td>Years employed</td>
<td>8.00</td>
<td>7.56</td>
</tr>
<tr>
<td>Number of cities resided in</td>
<td>5.84</td>
<td>4.22</td>
</tr>
<tr>
<td>Number of states/countries</td>
<td>3.00</td>
<td>3.28</td>
</tr>
<tr>
<td>Travel experience</td>
<td>7.74</td>
<td>5.67</td>
</tr>
<tr>
<td>Rural</td>
<td>.89</td>
<td>.72</td>
</tr>
<tr>
<td>Father's education</td>
<td>12.68</td>
<td>13.67</td>
</tr>
<tr>
<td>Mother's education</td>
<td>13.00</td>
<td>12.33</td>
</tr>
<tr>
<td>My education level</td>
<td>17.37</td>
<td>14.22</td>
</tr>
<tr>
<td>Helping profession</td>
<td>.63</td>
<td>1.00</td>
</tr>
</tbody>
</table>

**Note:** Asterisks indicate significant t-tests between Best and Worst means. Numbers in parentheses are n's.

a = 28 or greater on the Emotion Recognition Test.

b = 20 or less on the Emotion Recognition Test.

* p < .05 for a two-tailed test.

** p < .01 for a two-tailed test.

*** p < .001 for a two-tailed test.
One can see the comparison of scores for both the best and worst judges in Table 17. Only those variables having a statistically significant difference ($p < .05$) are employed to describe the high scorers. The result was that the following variables defined the best judges, in comparison with the worst judges: older, higher IQ, higher educational level, more flexible, more empathic, more traveled, and more likely a member of the helping professions. These results are quite similar to those found in Tables 3 and 6.

The best and worst judges on the stereotype other measure were compared in Table 18. As one can see from studying Table 18, there was only one significant difference between the high scoring and low scoring subjects on the measure of sensitivity to group differences. The good judge of the stereotype other tended to be older in age when he took his first job. With only one significant difference in 23 comparisons, the question is raised whether this difference was due to chance only.

Table 18

Means, Standard Deviations, and $t$-Values between the 17 Best$^a$ and 18 Worst$^b$ Judges on the Stereotype Accuracy Measure

<table>
<thead>
<tr>
<th>Variable</th>
<th>Best (17)</th>
<th>Worst (18)</th>
<th>Best (17)</th>
<th>Worst (18)</th>
<th>$t$-Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERT</td>
<td>24.59</td>
<td>22.89</td>
<td>3.94</td>
<td>5.22</td>
<td>1.08</td>
</tr>
<tr>
<td>IQ</td>
<td>103.35</td>
<td>117.44</td>
<td>29.37</td>
<td>20.81</td>
<td>1.65</td>
</tr>
<tr>
<td>Psychoticism</td>
<td>3.35</td>
<td>3.83</td>
<td>3.14</td>
<td>2.43</td>
<td>.51</td>
</tr>
<tr>
<td>Variable</td>
<td>Means</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-------</td>
<td>----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Best</td>
<td>Worst</td>
<td>Best</td>
<td>Worst</td>
<td>t-Values</td>
</tr>
<tr>
<td>Extraversion</td>
<td>13.53</td>
<td>13.38</td>
<td>4.43</td>
<td>4.90</td>
<td>.09</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>12.47</td>
<td>9.06</td>
<td>6.76</td>
<td>6.12</td>
<td>1.57</td>
</tr>
<tr>
<td>Lie Scale</td>
<td>4.59</td>
<td>4.17</td>
<td>1.97</td>
<td>2.64</td>
<td>.53</td>
</tr>
<tr>
<td>Psychological Mindedness</td>
<td>12.65</td>
<td>12.50</td>
<td>3.28</td>
<td>3.71</td>
<td>.12</td>
</tr>
<tr>
<td>Flexibility</td>
<td>12.65</td>
<td>13.22</td>
<td>4.00</td>
<td>4.57</td>
<td>.40</td>
</tr>
<tr>
<td>Hogan's Empathy Scale</td>
<td>25.59</td>
<td>25.44</td>
<td>5.05</td>
<td>5.12</td>
<td>.08</td>
</tr>
<tr>
<td>Stereotype Accuracy Measure</td>
<td>23.52</td>
<td>14.17</td>
<td>.61</td>
<td>1.29</td>
<td>26.61**</td>
</tr>
<tr>
<td>Sex</td>
<td>.53</td>
<td>.67</td>
<td>.51</td>
<td>.49</td>
<td>.81</td>
</tr>
<tr>
<td>Age</td>
<td>27.29</td>
<td>29.22</td>
<td>5.98</td>
<td>7.69</td>
<td>.82</td>
</tr>
<tr>
<td>Number of siblings</td>
<td>3.41</td>
<td>3.39</td>
<td>1.73</td>
<td>1.82</td>
<td>.03</td>
</tr>
<tr>
<td>Age of first job</td>
<td>17.47</td>
<td>15.72</td>
<td>2.50</td>
<td>1.71</td>
<td>2.43*</td>
</tr>
<tr>
<td>Number of different jobs</td>
<td>8.71</td>
<td>10.61</td>
<td>5.25</td>
<td>5.46</td>
<td>1.05</td>
</tr>
<tr>
<td>Years employed</td>
<td>7.88</td>
<td>10.06</td>
<td>5.95</td>
<td>7.50</td>
<td>.95</td>
</tr>
<tr>
<td>Number of cities resided in</td>
<td>4.71</td>
<td>6.00</td>
<td>2.76</td>
<td>3.11</td>
<td>1.30</td>
</tr>
<tr>
<td>during life time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of states resided in</td>
<td>2.65</td>
<td>3.83</td>
<td>1.73</td>
<td>1.95</td>
<td>1.90</td>
</tr>
<tr>
<td>during life time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Travel experience</td>
<td>6.88</td>
<td>8.00</td>
<td>2.39</td>
<td>2.61</td>
<td>1.32</td>
</tr>
<tr>
<td>Rural-urban background</td>
<td>.88</td>
<td>.78</td>
<td>.33</td>
<td>.43</td>
<td>.80</td>
</tr>
</tbody>
</table>
Table 18--Continued

<table>
<thead>
<tr>
<th>Variable</th>
<th>Means</th>
<th>SD</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(17)</td>
<td>(18)</td>
<td>(17)</td>
<td>(18)</td>
<td></td>
</tr>
<tr>
<td>Father's educational level</td>
<td>15.24</td>
<td>13.39</td>
<td>3.75</td>
<td>3.77</td>
<td>1.45</td>
</tr>
<tr>
<td>Mother's educational level</td>
<td>13.53</td>
<td>12.72</td>
<td>2.65</td>
<td>1.74</td>
<td>1.07</td>
</tr>
<tr>
<td>My educational level</td>
<td>16.41</td>
<td>16.89</td>
<td>3.41</td>
<td>2.99</td>
<td>.44</td>
</tr>
<tr>
<td>Helping profession</td>
<td>.71</td>
<td>.55</td>
<td>.47</td>
<td>.51</td>
<td>.90</td>
</tr>
</tbody>
</table>

Note: Asterisks indicate significant t-values between Best and Worst means. Numbers in parentheses are n's.

a = A good judge on the stereotype accuracy measure is one who scored 23 or greater on this test.
b = A poor judge on the stereotype accuracy measure is one who scored 15 or less on this test.

* $p < .05$ for a two-tailed test.
** $p < .001$ for a two-tailed test.

Summary of Results

The most capable judges of emotional expressions, in comparison with the least capable judges, were likely ($p < .05$) to have the following psychometric characteristics: higher IQ, more extraverted, more psychologically minded, more cognitively complex/flexible, and more empathic. The latter implies
effective functioning both socially and interpersonally. The capacity for independent thinking also characterizes the most capable judges. Experientially, the most competent judges were likely to be older, older in age at first job, and have fewer siblings. In contrast to the least able judges, the most competent ones had resided in more different cities, traveled more, experienced more formal education, and were more likely employed in the helping professions.

Fewer correlations were obtained between personality variables and a measure of stereotype accuracy, although three significant correlations were found. Being female, older in age at first job, and a higher scorer on Neuroticism were all characteristic of the most capable judges of stereotype accuracy.

Significant multiple linear regression equations were formed between the predictor variables and the two judgment tasks. Using six predictor variables, a multiple correlation (R) of .66 (R^2 = .44) was obtained for the Emotion Recognition Test. Two predictor variables yielded an R of .31 (R^2 = .09) for the stereotype accuracy measure. The success of several different predictors to estimate scores on the two judgment tasks suggested that the ability to judge others consists of many skills, not just one.

Dichotomizing the subjects into helpers and non helpers, the discovery was made that the helpers were significantly better on the Emotion Recognition Test, but no significant
differences were found on the stereotype accuracy measure. For those in the helping professions, years of experience in the helping professions were not related to success on either of the two judgment tasks.

Discussion

A discussion of the psychometric properties of the stereotype accuracy measure is important, as it may explain why so very few significant correlations were found with it. Of course, the possibility exists that there are but a few personality characteristics associated with being a good judge of stereotype accuracy, which is what the present study demonstrated; however, the conjecture is made that with more accurate measurement, additional characteristics of the good judge of stereotype accuracy will be identified.

The stereotype accuracy measure seemed to be less than desirable. The reason for the almost complete rejection of the hypotheses relating to the stereotype accuracy measure seemed to be because of its rather modest internal consistency (Cronbach's alpha = .60). In developing the stereotype accuracy measure, considerable care was given to the construction of it and the selection of items. The problem of low internal consistency seemed to be not only the length of the test, but the items contained in it did not discriminate between capable and incapable judges. Had the test
contained 100 items of similar discriminatory power, the internal consistency coefficient would have been well within acceptable limits (Cronbach's alpha = .86).

The methodological question was also raised about the use of items from personality inventories as part of a judgment task in person perception research. In the initial stages of developing the stereotype accuracy measure for this study, 123 MMPI items were employed. Examining the point biserial correlations for those items suggested that very few (less than 30) of those items discriminated between the most capable and least capable judges. With such a finding, the question is raised about the adequacy of internal consistency measures in other person perception tasks, which have used personality inventory items.

The hypothesis is that many researchers have assumed that because they were employing items from well-known personality tests, they had good internal consistency in their judgment tasks as well. In fact, they may have had a test, e.g., MMPI, that had good internal consistency for what it was designed to measure, but poor internal consistency as a judgment task. Hence, citing the internal consistency of judgment tasks seems important.

Despite the frustration of the stereotype accuracy measure, many of the findings of the study were quite exciting in their implications and ramifications for those interested in person
perception research. The noteworthy application of multivariate statistical procedures to person perception research is seen as the major contribution of this work. Only two known studies (Sharp, 1974; Zlatchin, 1973) have previously attempted to use multivariate procedures in person perception research, yet both had serious methodological defects (one study had too few subjects and too homogeneous a sample; the other, a criterion measure that had disastrously low statistical reliability). The present study is the first study to develop successfully multiple linear regression equations to predict measures of person perception. Much greater success was obtained in predicting those who are competent judges of sensitivity to individual differences than to group differences.

Despite the encouraging results that show that the use of multivariate procedures can assist in accounting for so much of the variance (44% of the variance of the Emotion Recognition Test), caution is still indicated. No cross-validation studies have been conducted with any of the multiple linear regression equations. Nonetheless, the application of multivariate research methods to the study of person perception seems encouraging for further research.

Related to the use of multivariate experimental procedures is the longstanding controversy over whether skill in person perception is one skill or many. Allport (1961) hypothesized that the ability to be a good judge of
others was neither entirely general nor entirely specific. This led to the first hypothesis: "There will be a significant positive correlation between the two types of person perception (stereotype accuracy and differential accuracy)." This hypothesis was supported; however, the amount of shared variance was so minuscule (less than 4%) that it hardly gives support to the notion of a general ability. Also, when one looks at the amount of variance (44%) accounted for in the Emotion Recognition Test by six different predictor variables, one has to question whether being a good judge of others is a general ability. Rather, the more plausible explanation is that numerous components are involved.

Previous research has been equivocal on this topic. Some research (Christensen, 1974; Cline & Richards, 1960; Sechrest & Jackson, 1961) has supported the view that a global ability to judge others exists. Other researchers (Bronfenbrenner, Harding, & Gallwey, 1958; Cline, 1964; Cronbach, 1955; Crow & Hammond, 1957) have stressed that numerous components exist in person perception, and these components are often independent from each other. Taguiri (1969) has argued that a person's accuracy in interpersonal perception cannot be described in terms of a single dimension or even a single component score. The present results would suggest otherwise. That is, through the application of multivariate experimental procedures, accuracy in person perception can be described in terms of a single score.
Just as the topic of whether empathy is one skill or many has a long history, so has concern over sex differences in person perception research. Hypotheses two and three both dealt with sex differences. For each hypothesis, no sex differences were predicted on either of the two judgment tasks. On the Emotion Recognition Test no sex differences were observed. This is consistent with Hoffman's conclusions (1977), but different from the results by Hall (1978) and Hall, Rosenthal, Archer, DiMatteo, and Rogers (1977). Hoffman did not find sex differences on recognition of affect, although he concluded sex differences existed on tasks measuring vicarious affective arousal. He based his information on a review of the literature; however, in 11 of the 12 studies he cited, the subjects were less than 11 years old!

The research by Hall and Hall et al. have suggested that females may be better than males at recognition of affect. These researchers found that females were better at judging nonverbal sensitivity than males at four different age levels (grade school, junior high, senior high school, and college). This study had a sample size over 2000, and it had better representation of the different age levels than the studies cited by Hoffman. Why the present study would have different results from Hall et al. is subject to much speculation. In contrast to either of those studies, the present one had older subjects ($\bar{X} = 27.9$, S.D. = 7.0) than Hall and Hall et al.
Also, there were many more different kinds of judgment tasks in the study by Hall et al. than in the present one.

One of the only significant correlations with stereotype accuracy measure was sex differences. Females were significantly better judges than males. This finding, of course, is consistent with much research in the person perception area; however, this is the first known finding which was specifically on a judgment task that was a measure of sensitivity to group differences. These results could reflect the "true" relationship between sex differences and stereotype accuracy, but they could also represent serendipity. Certainly more research will be necessary to clarify this relationship.

One of the most intriguing findings from the present research was that different personality and experiential characteristics for each sex were associated with being a good judge of others. Most often, variables that were important to one sex were also important to the other sex; however, there were notable exceptions. The reader is referred to Tables 3 and 6. Among the psychometric variables the observation was made that a low score on Psychoticism and a high score on Psychological Mindedness were associated with being a good judge on the Emotion Recognition Test for the female subjects, but not for the male subjects. Similarly,
low scores on the Lie scale were more associated with high scores on the Emotion Recognition Test among males, but not the females.

Significant sex trends were also noted among the experience variables. Age, age of first job, number of different cities resided in during lifetime, and travel experience were all correlated with accuracy of judgment for women, but not men. For males, fewer siblings, number of different jobs, and especially, educational level were all more associated with being sensitive to individual differences.

Hall et al. noted that the SAT math score correlated .30 with nonverbal sensitivity for males and .02 for females. Those authors also raised the question that the process of judgment may be somewhat different for males and females. They suggested that males may use a more analytical style; females, a more intuitive process. The present findings, together with those of Hall et al., suggest the following conundrum for researchers in the person perception area: Which variables, for which sex, are associated with being a good judge of others? Asking only which variables are associated with being a good judge of others will likely miss much.

Although research on sex differences in person perception has often been equivocal, research on intelligence has consistently been found to correlate with measures of person
perception (Taguiri, 1969). Hypotheses four and five predicted significant positive correlations between IQ and both measures of person perception. Hypothesis four was rejected, but Hypothesis five was accepted. There was a significant correlation \( r = .36 \) between IQ and a measure of sensitivity to individual differences. Along with Hogan's Empathy Scale, IQ had the highest correlation of the psychometric variables with the Emotion Recognition Test.

When significant positive correlations between IQ and person perception have not been found in past research, the reason has often been because of the experimental design. For example, Levy, Orr, and Rosenzweig (1960) had their subjects differentiate photographs into two groups: happiness vs. sadness. The subjects were college students and mental retardates. No relationship was found between IQ and the ability to make such differentiations. Data presented by Ekman and Friesen (1976) and Izard (1971) suggested that judging along the happiness dimension is the easiest of the facial expressions tasks. Hence, the results by Levy et al. may largely be an artifact of the particular stimulus object.

Sharp's finding (1974) that no significant relationship existed between intelligence and the ability to identify the affective state of another person pointed up one other serious flaw in trying to compare studies. All his subjects
were graduate students, suggesting such a homogeneous group intellectually that obtaining any correlation with IQ for this group would be difficult because of the restricted range of IQ. In summary, intelligence is obviously an important personality dimension with respect to explaining the ability to understand others, although its relevance seems to vary with the complexity of the judgment task and the degree of heterogeneity of the subjects.

Cognitive complexity/flexibility, a facet of intelligence, provided some interesting results. The reader is reminded that the construct cognitive complexity was measured by the Flexibility (Fx) scale of the California Psychological Inventory (CPI). Cognitive complexity/flexibility was studied in Hypotheses six and seven. The hypothesis was made that cognitive complexity/flexibility would be significantly positively correlated with both measures of person perception. Cognitive complexity was not correlated with the measure of stereotype accuracy; however, a significant correlation was found between Fx and the ability to identify emotions from facial expressions. Hjelle (1969) similarly found that his good judges of others had higher scores on Fx, although the differences did not reach statistical significance.

Just as cognitive complexity/flexibility represents a cognitive style, so does psychological mindedness. The relationship between Psychological Mindedness (Py) and
person perception was studied in Hypotheses eight and nine. In both hypotheses, the prediction was made that $Py$ would be significantly positively correlated with both measures of person perception. No significant relationship was found between $Py$ and measures of stereotype accuracy; however, a significant correlation was found between $Py$ and a measure of sensitivity to individual differences. Previous research (Burkhart, Gynther, & Christian, 1978; Hjelle, 1969; Wolitzky & Reuben, 1974) found similar trends. Burkhart et al. studied individuals' ability to discriminate between obvious and subtle MMPI items which measured psychopathology. Their results indicated that the high scorers on $Py$ were much better able to discriminate the pathological items (both obvious and subtle) than low scorers, suggesting that the high scorers are better able to sharply discriminate the social meaning of an interpersonal communication.

Both $Fx$ and $Py$ have been explored atomistically. Examining the underlying factors of those two scales adds to the conceptual knowledge of person perception, as both abilities are associated with being a competent judge of others. Examination of the factor analytic studies of the CPI reveals that the most frequently found third factor strongly loads on $Fx$, and occasionally $Py$. Megargee (1972) found that the most frequent meaning given to this Factor 3 is "capacity for independent thought and action" (p. 112).
Typically higher scorers on Factor 3 are independent thinkers who reject authoritarian attitudes. They tend to have higher IQ's, and they reject standardized, conventional, dogmatic solutions to problems.

While cognitive factors are important characteristics of the perceiver, the emotional state of the judge is equally as important; therefore, the effects of psychopathology on person perception were studied. Hypothesis 10 through 14 all predicted significant negative correlations between the two measures of person perception and the two measures of psychopathology (the Neuroticism and Psychoticism scales from the Eysenck Personality Questionnaire). None of these hypotheses was supported; nonetheless, important trends were noted, and they are discussed. Also, one of the ten groups included in the study consisted of psychiatric patients. Discussing the results gives the reader an additional opportunity to evaluate the effects of psychopathology on person perception.

The finding that Neuroticism was positively correlated with sensitivity to individual differences was quite surprising, because most other researchers (Bergin & Solomon, 1970; Vingoe & Antonoff, 1968; Hjelle, 1969) have found just the opposite relationship. One of the interesting paradoxes of this was that for those in the helping professions, Neuroticism
was significantly positively correlated \( (r = .48) \) with doing well on a measure of sensitivity to individual differences. Although neurotic characteristics are frequently thought of as negative, one of the positive characteristics is compulsivity. One of the correlates of compulsivity is being conscientious. Conceivably, the more conscientious, i.e., more neurotic subjects, put more effort into taking the tests. In moderation, compulsivity can be an asset to achievement. Professional students exemplify this!

The hypothesis between a measure of the stereotype accuracy and Neuroticism was predicted to be significant and negative. That hypothesis was rejected. The observation was made, however, that a significant positive correlation \( (r = .20) \) existed between the two variables. Again the question is raised: Did the "good effects" of neuroticism outweigh the bad effects?

The finding that Psychoticism was not significantly correlated \( (r = -.11, p > .05) \) with the Emotion Recognition Test was also surprising; however, closer inspection of Table 3 indicates that Psychoticism was significantly negatively correlated for the female subjects \( (r = -.24) \), but not the male subjects \( (r = .09) \). No significant relationship was observed between the stereotype accuracy measure and Psychoticism.

Why a stronger and more consistent relationship between Psychoticism and the Emotion Recognition Test was not found
seemed rather perplexing. After all, perceptual disorders constitute one of the cardinal symptoms of psychoses. Reading the definitive work on the Eysenck Personality Questionnaire by Eysenck and Eysenck (1976) suggests an explanation.

They noted that not everyone who scores high on Psychoticism is psychotic. Often, unconventional individuals will have elevated scores on that measure. In addition, two clinical groups tend to score high on Psychoticism: psychotics (especially schizophrenics) and psychopathic types (criminals). The latter characteristically score higher than the former! The two groups can frequently be distinguished from each other by looking at their respective scores on the Lie scale. Both psychotics and psychopaths have elevated Psychoticism scores, yet their Lie scores characteristically differ. Psychotics usually have high Lie scores; psychopaths, low scores. Despite the inclusion of 10 psychiatric patients (not all of whom had a psychotic diagnosis) in this study, the majority of the high scorers on Psychoticism in this study probably were not psychotic. Rather, they were more likely unconventional individuals. The reader may recall that Hekmat, Khajavi, and Mehryar (1975) found significant negative correlations between a measure of empathy and indicators of psychoses and neuroses on Lanyon's Psychological Screening Inventory; but no relationship was found between empathy and a measure of nonconformity (prisoners formed the normative group).
The question may still remain why Psychoticism was significantly negatively correlated for the female subjects, but not the males. The reason may be that a higher percentage of the female subjects had more severe psychological problems because of the sex ratio of the Psychiatric Patients' group. Most of those subjects (8 of the 10) comprising that group were females.

Despite the inclusion in the sample of 10 former psychiatric patients, the mean and standard deviation on Psychoticism were still quite small for the whole group ($\bar{X} = 3.14, \text{S.D.} = 2.87$). These figures are quite similar to those norms supplied by Eysenck and Eysenck (1975). A frequency count of the scores on Psychoticism for the 115 subjects revealed a very narrow range of scores for most of the subjects. Over 83% of the subjects scored between 0 and 5, and only 17% had scores of 6 or greater, suggesting a positively skewed distribution. The possibility exists that psychoticism is not related to person perception, which seems highly unlikely. More likely, the present measure of psychoticism, with its narrow range and skewed distribution, could have caused the non-significant correlation. That is, without some variability on a measure, obtaining significant correlations is very difficult.

One other measure of psychopathology is psychiatric patients themselves. Such a group was included in this
study. The reader is encouraged to examine Tables 7 and 8, and compare the performance of the Psychiatric Patients with the other groups. Observation of Table 8 indicates that the Psychiatric Patients had the very worst scores of the 10 groups on a measure of sensitivity to individual differences. This finding is consistent with other research (Dougherty, Bartlett, & Izard, 1974; Winkelmayer, Exline, Gottheil, & Paredes, 1978). Problems in accurate perception are most frequent among those suffering from various kinds of psychopathology. Hence, these results were expected.

Examination of Table 7 reveals that the Psychiatric Patients were the second best judges of all the groups on the stereotype other measure! Why this should be so is frankly baffling. Answering MMPI items like particular groups is not a visual perceptual task, as the Emotion Recognition Test is. Nonetheless, those with emotional problems typically have difficulty accurately interpreting social cues. Hence, the results are quite enigmatic. One should keep in mind, however, that looking at all the significant differences among the various groups indicated that the mean for the Psychiatric Patient group was significantly higher in comparison with only three of the other nine groups. It is interesting to note, however, that two of those groups were the PhD Clinicians and the Graduate Students in Psychology!
Extraversion, in contrast to the psychopathological dimensions of neuroticism and psychoticism, is more concerned with the social orientation of the individual. Hypotheses 14 and 15 focused on the relationship between Extraversion and measures of interpersonal judgment. In both cases, a negative correlation was predicted between Extraversion and the two judgment task measures. Neither hypothesis was confirmed, and no significant positive correlations were found between Extraversion and stereotype accuracy. Interestingly enough, however, a significant positive correlation was found between Extraversion and the ability to recognize facial expressions. Inspection of Table 3 indicates that this relationship was quite consistent, no matter which group of subjects was being studied.

Similar findings, using the Myers-Briggs Type Indicator as a measure of extraversion, were obtained in the landmark study done by Hall, Rosenthal, Archer, DiMatteo, and Rogers (1977). They found a positive relationship between extraversion and nonverbal sensitivity. Two studies by Hekmat, Khajavi, and Mehryar (1974, 1975) suggested that extraversion, not introversion, was more related to good empathic ability. While these studies suggest that extraversion is related to person perception, other researchers (Taft, 1950; Vernon, 1933; Vingoe & Antonoff, 1968) have presented data suggesting that introversion is more closely related to person perception.
than extraversion. Allport (1961) believed introversion was more associated with accurate person perception. One problem in comparing studies on extraversion is the surprisingly low correlations among the various measures of empathy (Cattell & Kline, 1977; Hundleby & Conner, 1968).

Other data from the present research suggested that being socially oriented, which is part of the definition of extraversion, is associated with being a good judge of others. The reader is referred to Table 8, which contains the scores on the Emotion Recognition Test. The top four groups on this measure were all from groups in Holland's typology that are social occupations. The four other occupational groups (based upon Holland's typology) all followed the social groups. The conjecture is made that those individuals who are oriented towards people, and hence those who spend more time around people, would have greater experience in judging people.

Reading Allport's view (1961) on the importance of introversion in person perception, one can see that he meant that the individual was capable of detachment. No doubt the capacity for detachment is related to introversion; however, equally plausible is the idea that the measure of extraversion misses too much. Cattell and Kline have forcefully argued this point. That is, one can imagine two individuals who are equally introverted, but some other personality
trait, e.g., dependency, would influence whether the individual was capable of detachment. The conjecture is made that the introverted but independent individual would be a better judge than the introverted but dependent individual. The latter might be so affiliative, pitying towards others, and insecure, that he could not take an impartial view of others. This conjecture awaits further research.

Although empathy, like extraversion, refers to social behavior, empathy is more concerned with social sensitivity, rather than just mere sociability. Hogan's Empathy Scale, the measure utilized to assess an empathic style, provided some of the most fruitful findings of the study. It was employed in both Hypotheses 16 and 17, which predicted that an empathic style would be significantly positively correlated with both measures of interpersonal perception. None of the psychometric variables tested in this study had higher correlations ($r = .38$) with sensitivity to individual differences than did Hogan's Empathy Scale. Looking back to Table 3, the reader can see that this relationship was sustained for almost all groups, even when IQ was partialed out. Certainly the present study provides additional validity to Hogan's Empathy Scale. Also, because this measure appears to be so consistently related to accuracy in interpersonal perception, further research with it would seem most worthwhile.
The groups that are assumed to be the most empathic are those in the helping professions, and yet the literature (Sarbin, Taft, & Bailey, 1960; Wiggins, 1973) comparing those in the helping professions with others on various social judgment tasks has not been encouraging. In general, research has not supported the view that those in the helping professions make the best judges of others. Examination of Table 8 offers some encouragement to those in the helping fields, and presents some interesting data to those interested in the psychology of vocational choice.

A comparison was made between the Helpers and the Non-helpers on the Emotion Recognition Test. The finding that those in the helping professions were found to be significantly better judges of affective expression than the non-helping group was more heartening. With more and more attention being given to the importance of nonverbal sensitivity in the helping process, this result is most reassuring.

Despite this optimism concerning those in the helping professions, there is also some worthwhile knowledge to be gained from observing their performance on the stereotype accuracy measure. Inspection of Table 7 reveals that two of the three helping groups (Experienced Clinicians and Graduate Students in Psychology) were the two worst groups on the stereotype other measure, whereas the third helping group (Counselor Education Students) was the best of the
10 groups on this measure. In spite of the overall poor performance of the helping groups on this task, there was no significant difference between the Helpers and the Non-helpers on this measure. Despite the lack of a statistical difference, the point is still clear: The Helpers did not perform especially well on this task. This points up one of the problems for those working in the helping professions. Some have called it the "clinician's error." This means, in essence, that the Helpers tend to see more psychopathology in people than really exists. For someone to obtain a low score on the stereotype accuracy measure, which two of the three helping groups did, they would have had to endorse items in the more pathological direction.

The very best judges on the Emotion Recognition Test came from the following four groups: Psychology Graduate Students, Experienced Clinicians, Education Students, and Counselor Education Students. All four of these occupations are oriented towards helping/working with people. In Holland's typology (1966), these are all Social types. The other four Holland occupational types in this study (Conventional, Artistic, Enterprising, and Intellectual) all scored less than those in the Social occupations. Such a finding seems consistent with other data presented by Holland, and adds to the body of knowledge of vocational choice.
Just as membership in the helping professions is assumed to be related to accuracy in judging others, years of experience in the helping professions, at least in the folklore, are typically associated with greater competence. For those in the helping professions, years of experience were not found to be related to success in either judging emotions from photographs or in guessing the right answers on a stereotype accuracy measure. These findings did not support Hypotheses 18 and 19, which stated that a significant positive correlation would be expected between years of experience in the helping professions and success on both the judgment tasks. This finding, unfortunately, is consistent with previous research (Goldberg, 1968; Sarbin, Taft, & Bailey, 1960; Wiggins, 1973).

The clinical versus statistical prediction controversy, spawned by Paul Meehl (1954), was discussed earlier in this paper, and seems germane to the present topic. After receiving numerous vilifying comments from other psychologists because of his alleged pro-statistical, anti-clinician bias, Meehl felt that he should clarify his views with his critics. The result was his article "What Can the Clinician Do Well?" (Meehl, 1967/1973). Meehl said, in brief, that one of the clinician's main assets is his ability to generate predictions. For example, the experienced clinician might think to himself: "I have seen others like this client...the extreme depression, the recent loss of a loved one, the rather
desultory look in the eyes...I should check into the possibility of suicide with him." The conjecture is made that with greater experience, the experienced clinician would generate more accurate predictions than his less experienced colleagues.

The point is that in the typical study in which the experienced clinician is pitted against his more inexperienced colleagues, the judgment task is characteristically highly structured ("answer these MMPI items like a 16 year-old male"), rather limited in scope ("place the MMPI's into either a neurotic or psychotic grouping"), and usually requires little creativity. No known studies exist which compare experienced clinicians with inexperienced clinicians when a hypothesis-generating task was the judgment task. Additional research will support or reject this conjecture that the experienced clinicians would be better than their less experienced co-workers at generating hypotheses.

Experience in its most general meaning, according to Allport (1961), was thought to be one of the most important characteristics of the good judge of others. Despite the fact that he identified experience as an important variable in being a good judge of others, little research has been done in this area. Numerous reasons exist for this. The present study was exploratory in that attempts were made
to investigate the relationships among measures of person perception and several experience measures. Those experience variables were the following: age, age of first job, number of siblings in the family of origin, number of different jobs in life time, number of years employed, number of states/countries resided in during life time, number of cities resided in during life time, travel experience, and educational level. Some of the more important findings related to those variables are discussed below.

Age has only briefly been researched. Taft's review of the literature (1955) indicated that developmental trends were apparent in their relationship to the ability to judge others; however, those trends seemed to stop at early adulthood. Hall, Rosenthal, Archer, DiMatteo, and Rogers (1977) found that nonverbal sensitivity increased linearly until "somewhere between twenty and thirty years of age" (p.202). Hence, two studies concluded that accuracy did not increase after early adulthood. Taft himself added, however, that further developmental trends in adulthood were difficult to assess because of two reasons: Most research had not employed subjects with a wide range of ages in the adult ranges, and the judgment tasks usually had methodological weaknesses ("answer the personality inventory items like a
college student"), which made them biased towards younger adults. Hence, the relationship between age and interpersonal accuracy is still unclear.

The present study revealed a significant positive correlation between age and the ability to recognize facial expressions. This was true even when the effect of having IQ partialed out was taken into consideration. This trend seems especially interesting when one considers that the general age of the subjects of this study tended to be older ($\bar{X} = 27.9$, S.D. = 7.0) than the usual study employing college students, who tend to average around 20 years of age.

Size of family has sparsely been researched, and the results have been mixed. Estes (1938) found no relationship between "sibling status" and the ability to judge others, whereas Taft (1950) found a negative correlation between the ability to judge others and number of siblings.

The present study was consistent with Taft's results: A significant negative correlation existed between number of siblings in the home and sensitivity to individual differences. Interestingly enough, this trend was more marked among men ($r = -0.46$) than women ($r = -0.28$). With fewer siblings, a child would have greater contact with his parents. Because adults are better judges than children, those from smaller families would have had greater exposure to the more interpersonally sensitive adult models, therefore, making them better judges.
Age of first job was thought to be related to an individual's experiences in life. Although not hypothesized, the guess was made that one who went to work earlier in life would have had a greater variety of experiences with others; therefore, he would be a better judge of others. Just the opposite trend was found for both judgment tasks. Those who went to work later tended to be better judges on both judgment tasks. Precisely what this means is open to much speculation. The observation was previously made that age of first job was significantly correlated with both father's educational level \( r = .25 \) and mother's educational level \( r = .18 \). As educational level is positively correlated with social class, the hypothesis is made that age of first job may be a distant measure of social class. Interestingly enough, the relationship between age of first job and scores on the Emotion Recognition Test was even stronger \( r = .30 \) when IQ was partialed out.

No known studies exist that relate age of first job to accuracy of interpersonal judgment; however, two known studies have evaluated the relationship between social class and person perception. Sweet (1929) found a positive relationship between socioeconomic status and empathic ability, but Taft (1950) found no such relationship among graduate students. The latter subjects were all graduate students, suggesting a very restricted range on a socioeconomic measure.
Travel experience was assessed by means of three different measures: number of cities resided in during life time, number of states/countries resided in during life time, and a self-evaluation of travel experience. The supposition was made that an individual who traveled more would have had a wider range of experiences. Significant correlations were found for two of the three variables with the Emotion Recognition Test, although no significant correlations were found for number of states/countries resided in during life time. In addition, no significant correlations were found between the stereotype accuracy measure and any of the travel variables.

While correlations for all 115 subjects for two of the variables (travel experience and number of cities resided in during life time) were found, they were much higher for the female subjects than the male subjects, suggesting that such experience was more important for females. The reader is referred to Table 6 to see this trend more clearly. For the female subjects the correlations between travel experience and the ability to recognize facial expressions was one of the highest found in the study ($r = .41$).

Educational level was also investigated. Typically, this variable has not been explored, except in the comparison of individuals (usually children) of different age levels. Among adults, no known studies exist that compared educational level with the ability to judge others. The present findings
indicated that educational level was very significantly related ($r = .36$) to being a good judge of facial expressions. This relationship was sustained, even when IQ was partialed out. Also of some significance is the finding that educational level was more important to males ($r = .51$) than females ($r = .33$). This result is not considered unusual, when one considers that males typically become more sensitive and have greater aesthetic values as a result of a college education (Feldman & Newcomb, 1969).

**Limitations**

This study had several limitations. Most importantly, the subjects were volunteers. This could have influenced their response style in some way. The assumptions were made that all the subjects cooperated, gave their best effort, and responded honestly to the various questionnaires.

Results from the Eysenck Personality Questionnaire indicated that the subjects tended to answer that test honestly, which may imply that they answered the other personality measures genuinely. This was assessed by two statistics. Both the mean score on the Lie scale, which was within normal limits, and the low correlation ($r = -.04$) between the Lie scale and the Neuroticism scale insinuate an honest response style. If the latter two scales have a high negative correlation between them, it suggests that the subjects were "faking good" (Eysenck & Eysenck, 1976). Obviously that was not the case in the present study.
The subjects did not represent an accurate cross section of the American public, although the sample was more heterogeneous than most samples employed in person perception studies. This study had the usual limitation of using college and college-educated subjects. Hence, caution is advised in generalizing these results beyond the present population.

One limitation which was recognized in using self-report personality instruments is that subjects only reveal what they choose. Also, no correction for various response styles ("faking bad," "faking good," or random marking) was available. The statistical reliabilities of two of the personality scales (Flexibility and Psychological Mindedness) were less than one would desire for good measures of each; however, there simply are no well-known established measures of each.

The judgment tasks used in the present study also had certain limitations. The photographs employed for the Emotion Recognition Test were exclusively of Caucasians. The measure of the stereotype accuracy was developed for purposes of this study. Its internal consistency was less than desired (Cronbach's alpha = .60), and it proved too short, which made its range too narrow. While both measures of person perception (differential accuracy and stereotype accuracy) have been utilized, one cannot say that these two judgment tasks encompass all facets of person perception.
Conclusions

The following conclusions, based on this study, were made:

1. Individuals who are likely to be the most astute judges of others, in comparison with the least astute judges of others, are likely to be intelligent individuals who function effectively in social and interpersonal situations. Furthermore, they are mature individuals who have a wide range of experiences in the world. They are characteristically independent thinkers who reject dogmatic solutions. As one can see, this description of the most capable judges of others is not that different from a typical description of the mentally healthy individual.

2. The success of the multiple linear regression equations to predict so much of the variance of the Emotion Recognition Test certainly argues that being a good judge of others is made up of many skills, not just one.

3. For those in the helping professions, years of experience in the helping professions do not appear related to success in person perception, at least when the judgment tasks are highly structured.

4. For each sex, different personality characteristics and different kinds of experiences are associated with being a competent judge of others.
5. General experiences in life such as maturity, travel experience, and education, appear to be related to accuracy in person perception.

6. The present data supports the validity of Hogan's Empathy Scale.

7. Individuals who are, in Holland's typology of vocational choice, Social types, i.e., those engaged in occupations oriented towards helping others, appear to be the most accurate judges of others.

Recommendations

On the basis of the findings of this investigation, the following recommendations are made:

1. Despite the numerous measures taken in the present study, there was simply a limit to how many tests subjects would have willingly taken. Nonetheless, additional measures might have improved the study. A measure of femininity is suggested, because those who are more feminine (both males and females) may very well be better judges. Jung's discussion of the anima and animus would suggest this.

Although a measure of cognitive complexity was made by using the Flexibility scale from the California Psychological Inventory, a more valid measure would have been the Cognitive Complexity scale from the recently developed Jackson Personality Inventory. Kelly's Rep Test could also be employed, but it would take considerably longer than the 20-item Cognitive Complexity scale.
Allport (1961) identified artistic interest as a correlate of accuracy in person perception. A future study in person perception might include the Barron-Welsh Art Scale. Allport also suggested that insight was important to person perception; therefore, future studies might include such a measure.

The Chapin Social Insight Test offers an interesting possibility as a measure of person perception. In brief, it is a 25-item test that requires the reader to understand a situation involving human interaction, recognize the underlying personality dynamics of the characters involved, and choose the best method for the resolution of the conflict between the characters.

The research by Wolfe has very much shown the importance of social desirability and assumed similarity as components in person perception. Not only has he shown the importance of these measures, but he has provided a model to measure these and utilize them in research.

2. The Emotion Recognition Test is suggested as a judgment task for further research. Not only did it have good internal consistency (Cronbach's alpha = .71), but the affective response of the subjects to it was quite favorable.

3. Separation of the sexes in analyzing data in person perception seems most important. Despite the fact that the same personality and experiential variables were often associated with accuracy in person perception for both sexes,
the observation was made that different characteristics were often associated with each sex doing well on a person perception task.

4. Measures of differential accuracy and stereotype accuracy are needed in any study of person perception. The correlations of success on both tasks were quite different.

5. To more fully explore the correlates of stereotype accuracy, a more statistically reliable measure is needed than the present one.
Appendix A

FORM A

USE OF HUMAN SUBJECTS

COMMITTEE ACTION

1. Activity Title: A Study of Person Perception Based on Selected Personality Variables Identified by Gordon Atiport
2. Activity Director: Kay Hale
3. Department: Education
4. Telephone Extension: 2736
5. Date Submitted: 10 April 1973

X The statement submitted for the activity conforms to the University Policy on the Protection of Human Subjects and the activity is approved.

The statement submitted for the activity does not conform to the University Policy on the Protection of Human Subjects and the activity is disapproved for reasons stated on the attached sheet(s).

Signature: Robert W. Grady, Chairman
Institutional Review Board for the Protection of Human Subjects

April 28, 1973
In response to your request of March 10, 1973, permission is hereby granted you to reproduce the items on the Anxiety, Psychological Hardiness, and Flexibility scale of the CPI for use in your Ph.D. dissertation research. This permission is restricted to no more than 150 copies, subject to the following restrictions:

(a) Any and all material used will contain proper acknowledgments; e.g., “Reproduced by special permission from the California Psychological Inventory by Harrison Gough. Copyright date: 1958. Published by Consulting Psychologists Press Inc.”

(b) None of the materials may be sold or used for purposes other than those mentioned above.

(c) One copy of any material reproduced will be sent to the Publisher.

CONSULTING PSYCHOLOGISTS PRESS INC.

By [Signature]
Permissions Editor
Date [Handwritten Date]
Appendix C

To the Subjects:

PURPOSE OF THIS RESEARCH:
How we come to understand others (empathy) is an important skill for everyone. Despite the significance of this topic, much is still not understood about this process. The purpose of this research is to clarify some of the personality characteristics that are associated with understanding others.

The following tests will be administered to you:
1. an intelligence test
2. two different personality tests
3. thirty-six photographs will be shown to you and you will be requested to identify the emotions being expressed.
4. you will be asked to answer 34 questions from a well-known personality inventory like you think most 16 year olds and most 70 year olds would.
5. a personal data sheet will request certain personal and demographic information from you.

I will answer any questions you have about the procedures for this test. You are free to withdraw your consent and to discontinue participation in the project at any time.

ANONYMITY:
No attempt will be made to identify any subject.
Do not put your name on any answer sheet or test booklet, although I will ask you to put the last four digits of your social security number on the answer sheets in order to keep track of each subject's test results.

INFORMED CONSENT:
Name of subject:________________________
1. I hereby give consent to Kay Hale to perform the above described experimental procedure.
2. I have heard a clear explanation and understand the nature and purpose of the procedure. I have heard a clear explanation and understand the benefits to be expected. I understand that the procedure to be performed is investigational and that I may withdraw my consent for my participation at any time. With my understanding of this, having received this information and satisfactory answers to the questions I have asked, I voluntarily consent to the procedure described in the first paragraph.

________________________
Date

________________________
Signature of Subject

REPRINT REQUEST:
All too often people take part in research but never hear what the results were. When my dissertation is completed (Dec., '78), I will write a several page summary of the results. If you would like a summary of the findings, put your name and address below.

Thank you for your participation.

Name:________________________________
Address:______________________________
City:______________________

Sincerely,

Kay Hale
Appendix D

Personal Data Sheet

Instructions: Please answer all questions. Guesses are better than no answers. Mark all answers on the Personal Data Answer Sheet.

1. Sex: 0 = Female
   1 = Male

2. I am: 1. Caucasian
   2. Black
   3. Mexican-American
   4. Oriental
   5. Other/Prefer not to say

3. Age:

4. The number of children in your family while growing up (count yourself):

5. How many older brothers do you have?

6. How many younger brothers do you have?

7. How many older sisters do you have?

8. How many younger sisters do you have?

9. Age of your first job which required you to work at least 20 hours every week. (Do not count any jobs related to household duties or chores.)

10. How many different jobs (both full-time and part-time) have you had in your life time?

11. For how many years have you been employed in your life time?

12. How many different cities/towns have you resided in during your life time?

13. How many different states/countries have you resided in during your life time?

14. Compared to others my age, I have traveled:

   Very Much
   Less
   Than Most
   1 2 3 4 5 6 7 8 9 10 11
   An
   Average
   Amount

15. Did you grow up primarily in which of the following:
   0 = rural area
   1 = urban area
16. Highest grade level completed by your father:  
Use the following as a guide:  
ninth grade = 9  
high school graduate = 12  
one year of college = 13  
college graduate = 16  
master's degree = 18  
law degree = 19  
PhD/EdD/M.D. = 21  

17. Highest grade level completed by your mother:  
Use the above as a guide.  

18. Highest grade level completed by you:  
Use the above as a guide.  
If you are currently in school, count through last semester as your  
highest grade completed.  

19. Are you currently in school?  
0 = yes  
1 = no  

20. If you are in school, what is your classification? (If not in school  
currently, write down a 7).  
0 = High school student  
1 = Freshman in college  
2 = Sophomore in college  
3 = Junior in college  
4 = Senior in college  
5 = I'm working on a master's degree  
6 = I'm working on a doctorate  

21. I work in a helping profession or I am studying to be a helping  
professional (Helping profession is defined here as psychologist,  
counselor, social worker, or behavior modification specialist.)  
0 = yes  
1 = no  

22. What is your area of employment or your major in school?  
1 = Psychology  
2 = Counselor Education  
3 = Fine arts (either art, music or drama)  
4 = Biology/Chemistry/Physics  
5 = Education  
6 = Business  
7 = Secretarial/clerical worker  
8 = Other (Please write out specifically)  

23. My current occupation or title of my job is: (write out on answer sheet).
24. How many years of actual experience have you had in the helping profession? Guide: (Helping profession is defined here as psychologist, counselor, social worker, or behavior modification specialist.) "Experience" in the helping profession is limited to experiences doing only the following: (a) counseling/psychotherapy, (b) diagnostic work, (c) behavior modification programs. It would not include such activities as camp counselor or dorm counselor.

1. If you are not in the helping profession, place a 0.
2. One year in the helping profession is defined here as 40 hours per week for approximately 50 weeks per year.
3. Count practicums and internships.
4. Example: two years of working twenty hours per week would equal one year in the helping profession.
Appendix E

Last 4 digits of your SS #

Personal Data Answer Sheet

Instructions: Unless asked for, answer all questions by number, rather than words, e.g., on question 1, you are asked if you are a female = 0; male = 1. Write down either a 0 or 1, do not write down male or female.

1._____
2._____
3._____
4._____
5._____
6._____
7._____
8._____
9._____
10._____
11._____
12._____
13._____
14._____
15._____
16._____
17._____
18._____
19._____
20._____
21._____
22._____
23.__________________________
24._____


Appendix F

Personality Test

1. A person needs to "show off" a little now and then.
2. I liked "Alice in Wonderland" by Lewis Carroll.
3. People can pretty easily change me even though I thought that my mind was already made up on a subject.
4. I often feel that I made a wrong choice in my occupation.
5. I always like to keep my things neat and tidy and in good order.
6. Clever, sarcastic people make me feel very uncomfortable.
7. I usually take an active part in the entertainment at parties.
8. I find it hard to keep my mind on a task or job.
9. I feel sure that there is only one true religion.
10. I am afraid of deep water.
11. I must admit I often try to get my own way regardless of what others may want.
12. I have at one time or another in my life tried my hand at writing poetry.
13. I don't like to undertake any project unless I have a pretty good idea as to how it will turn out.
14. Most of the arguments or quarrels I get into are over matters of principle.
15. I do not like to see people carelessly dressed.
16. The idea of doing research appeals to me.
17. I would like the job of a foreign correspondent for a newspaper.
18. People today have forgotten how to feel properly ashamed of themselves.
19. I cannot keep my mind on one thing.
20. I prefer a shower to a bathtub.
21. I believe we are made better by the trials and hardships of life.
22. I always try to consider the other fellow's feelings before I do something.
23. One of my aims in life is to accomplish something that would make my mother proud of me.
24. I have a tendency to give up easily when I meet difficult problems.
25. A person should adapt his ideas and his behavior to the group that happens to be with him at the time.
26. In school I always looked far ahead in planning what courses to take.
27. I do not have a great fear of snakes.
28. I often do whatever makes me feel cheerful here and now, even at the cost of some distant goal.
29. I usually don't like to talk much unless I am with people I know very well.
30. I can remember "playing sick" to get out of something.
31. I like to keep people guessing what I'm going to do next.
32. Before I do something I try to consider how my friends will react to it.
33. Sometimes at elections I vote for men about whom I know very little.
34. I have frequently found myself, when alone, pondering such abstract problems as free will, evil, etc.
35. I would like to write a technical book.
36. Most people worry too much about sex.
37. I much prefer symmetry to asymmetry.
38. We ought to pay our elected officials better than we do.
39. I like to talk before groups of people.
40. When a man is with a woman he is usually thinking about things related to her sex.
41. Only a fool would try to change our American way of life.
42. My parents were always very strict and stern with me.
43. Sometimes I rather enjoy going against the rules and doing things I'm not supposed to.
44. A large number of people are guilty of bad sexual conduct.
45. I think I would like to belong to a singing club.
46. I often wish people would be more definite about things.
47. It is annoying to listen to a lecturer who cannot seem to make up his mind as to what he really believes.
48. I find that a well-ordered mode of life with regular hours is congenial to my temperament.
49. It is hard for me to sympathize with someone who is always doubting and unsure about things.
50. I often start things I never finish.
51. Our thinking would be a lot better off if we would just forget about words like "probably," "approximately," and "perhaps."
52. I never make judgments about people until I am sure of the facts.
53. A strong person will be able to make up his mind even on the most difficult questions.
54. For most questions there is just one right answer, once a person is able to get all the facts.
55. I think I am usually a leader in my group.
56. I like to have a place for everything and everything in its place.
57. I don't like to work on a problem unless there is the possibility of coming out with a clear-cut and unambiguous answer.
58. It bothers me when something unexpected interrupts my daily routine.
59. Most of the arguments or quarrels I get into are over matters of principle.
60. I am known as a hard and steady worker.
61. I don't like things to be uncertain and unpredictable.
62. Once I have my mind made up I seldom change it.
63. I think I am stricter about right and wrong than most people.
64. I have a natural talent for influencing people.
65. I am in favor of a very strict enforcement of all laws, no matter what the consequences.
66. I always see to it that my work is carefully planned and organized.
67. I don't really care whether people like me or dislike me.
68. The trouble with many people is that they don't take things seriously enough.
69. I set a high standard for myself and I feel others should do the same.
70. People who seem unsure and uncertain about things make me feel uncomfortable.
71. It is hard for me just to sit still and relax.
72. As a rule I have little difficulty in "putting myself into other people's shoes."
73. I have seen some things so sad that I almost felt like crying.
74. Disobedience to the government is never justified.
75. It is the duty of a citizen to support his country, right or wrong.
76. I am usually rather short-tempered with people who come around and bother me with foolish questions.
77. I have a pretty clear idea of what I would try to impart to my students if I were a teacher.
78. I enjoy the company of strong-willed people.
79. I frequently undertake more than I can accomplish.
Emotion Recognition Experiment

Instructions: In a few minutes you will be shown 36 photographs. When you are shown photograph number 1, look first at the photograph, then at the list of emotions provided you. The list contains nine different emotions, lettered A to I. Select the one emotion that best describes the photo, then write the letter of the emotion opposite the number of the picture.

First take a few minutes to study the names and definitions of the Fundamental Emotions so you will be familiar with all of them.

A. INTEREST-EXCITEMENT: Concentrating, attending, attracted, curious.

B. ENJOYMENT-JOY: Glad, merry, delighted, joyful.

C. SURPRISE-STARTLE: Sudden reaction to something unexpected, astonished.

D. DISTRESS-ANGUISH: Sad, unhappy, miserable, feels like crying.

E. DISGUST-REVULSION: Repugnance, aversion, distaste, sickened.

F. ANGER-RAGE: Angry, hostile, furious, enraged.

G. SHAME-HUMILIATION: Embarrassed, ashamed, guilty, shy.

H. FEAR-TERROR: Scared, afraid, terrified, panicked.

I. CONTEMPT-SCORN: Disdainful, sneering, derisive, haughty.

*1. _____  13. _____  25. _____

*2. _____  14. _____  26. _____

3. _____  15. _____  *27. _____

4. _____  16. _____  28. _____

5. _____  17. _____  29. _____

6. _____  18. _____  30. _____

7. _____  19. _____  31. _____

8. _____  20. _____  32. _____

*9. _____  21. _____  33. _____

10. _____  22. _____  34. _____

11. _____  23. _____  35. _____

12. _____  24. _____  36. _____

* Omitted in the final data analysis.
Appendix H
Empathy Test

Answer all the below questions either TRUE (if it is true or mostly true) or FALSE (if it is false or mostly false). True = 1
False = 2

Answer these questions the way you think most 16-19 year old men would answer them:
1. I have a good appetite. (2, True)
*2. My speech is the same as always (not faster or slower, or slurring; no hoarseness.) (119, True)
*3. I am worried about sex matters. (179, False)
*4. I have been quite independent and free from family rule. (235, False)
*5. I often must sleep over a matter before I decide what to do. (402, False)
*6. I have often met people who were supposed to be experts who were no better than I. (406, False)

Answer these questions the way you think most 70+ year old women would answer them:
7. I think a great many people exaggerate their misfortunes in order to gain the sympathy and help of others. (71, True)
8. I like collecting flowers or growing house plants. (132, True)
9. I like to cook. (9, True)
10. I would like to be a journalist. (204, False)
11. I am about as able to work as I ever was. (9, False)
12. I do not like to see women smoke. (378, True)
13. I love to go to dances. (391, False)
14. A windstorm terrifies me. (392, True)
15. I have often met people who were supposed to be experts who were no better than I. (406, True)
16. I am attracted to members of the opposite sex. (430, False)
17. I have strong political opinions. (432, True)
18. Sexual things disgust me. (470, True)
19. I am greatly bothered by forgetting where I put things. (560, True)

Answer these questions the way you think most 70+ year old men would answer them:
20. I would like to be a singer. (20, False)
*21. At times my thoughts have raced ahead faster than I could speak them. (134, True)
22. I have been inspired to a program of life based on duty which I have since carefully followed. (232, True)
23. It takes a lot of argument to convince most people of the truth. (89, True)
24. What others think of me does not bother me. (170, True)
*25. I have strong political opinions. (432, True)
26. I have been inspired to a program of life based on duty which I have since carefully followed. (232, True)
27. I am entirely self-confident. (264, True)

Answer these questions the way you think most 16-19 year old women would answer them:
28. I have often wished I were a girl. (Or if you are a girl) I have been sorry that I am a girl. (74, True)
*29. I feel that it is certainly best to keep my mouth shut when I'm in trouble. (26, False)
30. I daydream very little. (30, False)
31. I have often felt that strangers were looking at me critically. (278, True)
32. Once in a while I feel hate towards members of my family whom I usually love. (282, True)
33. I dream frequently. (425, True)
34. I like to read newspaper editorials. (428, False)

* Omitted in the final data analysis.
References


Dickson, W.J. (Ed.) *Biomedical computer programs.* Los Angeles: Health Sciences Computing Facility, University of California at Los Angeles, 1970.


Gough, H.G. *The development of a rigidity scale.* Unpublished mimeographed manuscript, Institute of Personality Assessment and Research, Berkeley, California, 1951.


Gough, H.G. College attendance among high aptitude students as predicted by the California Psychological Inventory. *Journal of Counseling Psychology, 1968b, 15,* 269-278.


