A TEST OF STEREOTYPE ACCURACY

Thesis for the Degree of M. A.
MICHIGAN STATE UNIVERSITY
Albert Zavala
1960

097
P

174

S3 —— 069

I —— 101

V —— 127

S —— 751

26 —— 063

23
R
679

A TEST OF STEREOTYPE ACCURACY

by

Albert Zavala

Approved

ABSTRACT

A TEST OF STEREOTYPE ACCURACY

by Albert Zavala

The purpose of this study was to develop an objective and
economical test that would measure Stereotype Accuracy, or the ability
to predict the behavior of social groups.

Four 30-item scales were designed to measure the accuracy
of four stereotypes: College Man Stereotype, Woman Stereotype, Sex
Stereotype, and Age Stereotype. These were developed from four
preliminary subscales, using item analysis of the longer form. The
items in the test were multiple-choice items, with elements being
drawn from the Strong Vocational Interest Blank. The correct answer
for each item was that listed by Strong as having the highest preference
by persons of a given category (stereotype). Repeat reliabilities of
the stereotype scales and total scores ranged from . 37 to . 60. The
internal consistency, as measured by the Kuder-Richardson formula
#20, ranged from . 16 to . 58.

Thirteen additional measures were obtained from one group
of 100 students to explore the relationships of the measures to the
scores on the several scales. A total of 266 college students served
as subjects in the various phases of the study.

Evidence was found to support the following conclusions. The accuracy of a person's stereotype is related to his accuracy in predicting the behavior of individuals (r of .20 in one test, and r of .29 in another). The accuracy of a person's stereotype has little relation to his knowledge of psychology (r of .14 with course grade in psychology). The accuracy of a person's stereotype is related to his attitudes towards methods of successful leadership (correlations range from -.24 to .26).

As measured by the present instrument, stereotype accuracy appeared to be a highly specific ability, for the correlations between the subscales ranged from -.19 to .28. Whether this is an artifact of the measuring instrument and whether instruments with greater generality can be developed, are questions that can be resolved only by further experimental inquiry.

A TEST OF STEREOTYPE ACCURACY

by

Albert Zavala

A THESIS

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

MASTER OF ARTS

Department of Psychology

1960

Este tesis se la dedico a mi madre con todo

mi cariño. Si no fuera por ella, esta obra no

conoce la lúz del día.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

The purpose of the present study was to develop a test of stereotype accuracy, determine its homogeneity, internal consistency, and stability, and correlate it with variables which might provide some evidence of its validity. Stereotype accuracy, following the definition of Cronbach (1955), is the ability to predict the typical responses of a social group. Recent theoretical considerations and research results indicate that a person's stereotype accuracy is of central importance in contributing to the adequacy of his "empathy, " "interpersonal perception, " "clinical intuition, " and "sensitivity to others, " to mention a few of the concepts closely related to stereotype accuracy.

HISTORY

The history of the problem of stereotype accuracy starts with the concept of empathy. The term was introduced by Lipps in 1907. Through the implicit or stated definitions of many writers (Freud, as quoted in Allport, 1937; Baumgarten, 1931; Sullivan, 1940; Fleiss, 1942; Reusch, 1948; Hutt and Miller, 1949; Fromm-Reichmann, 1950; Rogers, 1950), the definition of empathy has developed to the form found in English and English (1958): "1 . . . attribution of the feelings or attributes aroused by its surroundings (actual or depicted) to a natural object or work of art . . . . 2. apprehension of the state of mind of another person without feeling (as in sympathy) what the other feels."

Many psychologists have been concerned with theoretical explanations of empathy (Luszki, 1951; Baxter, 1952; Halpern, 1954; Norman and Leiding, 1956; Vernon and Stewart, 1956). Both clinical psychologists (Escalona, 1945; Reusch and Prestwood, 1949; McClelland, 1951; and Cowden, 1955), and industrial psychologists (Kerr, 1951; Speroff, 1953, 1954; and Bell and Stolper, 1955) have viewed empathic ability as being of considerable importance in the areas of their interests. However, the analysis of experimental results indicated that the measurement of empathy and the interpretation of the measurements is peculiarly complicated.

Bender and Hastorf (1953) for example, found that a judge's predictions were much more highly related to his own responses than they were to the responses of the person whose behavior he was predicting. Thus, if it happened that the judge and the person being judged were actually similar, this would improve the empathy scores of the judge. In order to circumvent this problem, Bender and Hastorf developed a "refined empathy" score by subtracting the judge's projection score from his prediction score. The projection score in this case was the deviation of a predictor's own answers from his predictions for his associate.

Norman and Leiding (1956) investigated the relationships between the variables empathy and projection, and refined empathy with regard to individual and Mass Empathy tests. The Mass Empathy test measured a person's ability to predict the typical self-responses of a group of persons. This was called "raw empathy." The number of times his judgments for others' responses are the same as his own responses are totalled and called "projection." "Refined empathy" is taken as "raw empathy" minus "projection." Murstein (1957) gives a detailed account of the weaknesses inherent in the measures used by Norman and Leiding.

Cronbach (1955), in a more systematic and detailed analysis of processes affecting scores on the understanding of others, concludes that empathy scores are the sum of four components: Elevation, Differential Elevation, Stereotype Accuracy, and Differential Accuracy.

He defines these components as:

Elevation (E). The elevation component has the form

$(\bar{y}.._j - \bar{x}..)^2$, $\bar{y}.._j$ is the average of J's predictions over all items and

all O's; it reflects his way of using the response scale. The elevation

component is increased by any difference between J's central tendency

of responding and the central tendency of the self-descriptions, for

all items and O's combined.

Differential Elevation (DE). Differential elevation reflects

how closely J's average prediction for O corresponds to O's central

tendency of response, all items pooled and J's central tendency of

response held constant. That is, it reports J's ability to judge deviations

of the individual's elevation from the average. We may write DE in

this form:

$$DE_j^2 = \sigma_{\bar{y}_{o \cdot j}} + \sigma_{\bar{x}_o} - 2\sigma_{\bar{x}_{o \cdot}} r_{\bar{x}_{o \cdot} \bar{y}_{o \cdot j}} \qquad [1]$$

The variance $\sigma_{\bar{y}_{o \cdot j}}^2$ expresses J's report of how much O's will differ

in elevation. This assumed dispersion in elevation will appear later

as a component of the assumed similarity score. $\sigma_{\bar{x}_{o \cdot}}^2$ is the true

dispersion in elevation. The correlation $r_{\bar{x}_{o \cdot} \bar{y}_{o \cdot j}}$ (to be symbolized

DE_r) represents J's ability to judge which O's rate highest on the

elevation scale.

In some tests, central tendency of response (elevation)

reflects insignificant response sets. In other tests, elevation has an

important psychological meaning. Thus, if a high score on each item

indicates Morale, the correlation DE_r shows how will J can judge which O's say they have the highest Morale.

"<u>Stereotype Accuracy</u> (SA). Stereotype accuracy describes J's ability to predict the norm for Os. It might be called 'accuracy in predicting the generalized other.' This score depends on J's knowledge of the relative frequency or popularity of the possible responses.

"We may write:

$$SA_{.i}^2 = \sigma_{\overline{y}_{.ij}}^2 + \sigma_{\overline{x}_{.i}}^2 - 2\sigma_{\overline{y}_{.ij}} \sigma_{\overline{x}_{.i}} r_{\overline{y}_{.ij}\overline{x}_{.i}} \qquad [2]$$

Here each variance is computed over items. The variance $\sigma_{\overline{x}_{.i}}^2$ is the scatter of the actual means. SA represents ability both as to shape and scatter. $r_{\overline{y}_{.ij}\overline{x}_{.i}}$ (Stereotype Correlation, SA_r) represents accuracy in judging mean profile shape, i.e. the order of item difficulties.

"<u>Differential Accuracy</u> (DA). Differential Accuracy reflects ability to predict differences between Os on any item. This component is averaged over items . . . this component . . . may be regarded as combining variances with a correlation term. The correlation (DA_r) measures the ability to judge which Os have the highest scores on the item, when the score is taken as a deviation from Os mean. There is one such correlation for each item" (Cronbach, 1955).

Cronbach states that the failure to separate the components of empathy makes interpretation ambiguous. Where errors of

measurement affect one element or component of a global score, they

influence the others also. Thus "significance tests are spurious and

correlations are artifactually raised or lowered" (Cronbach, in Taguiri

and Petrullo, 1958). Two of Cronbach's examples serve to clarify

these points:

> One investigator hypothesized that the extent to which
> teachers knew their pupils' opinions would correlate with their
> effectiveness in establishing a good classroom atmosphere.
> He asked pupils to rate their teacher on patience, friendliness,
> etc., and used the rating, averaged over pupils as a criterion.
> Call this **X**. He also asked the teacher to predict what average
> rating he would be given (Y), and used D, the absolute difference
> between X and Y, as a measure to support the hypothesis that
> superior teachers have greater insight--but this result occurred
> simply because all teachers predicted a favorable Y for them-
> selves. Predictions varied so little than in effect D differed
> from Y only by a constant. If Y had varied substantially, there
> would still have been some artifactual constraint on the correlation.
> Linkages often arise in a disguise which makes them
> difficult to trace. One example is the work of Hastorf and Bender.
> They were concerned with "empathy," "projection," and "real
> similarity," as derived from self-perception of another person,
> and their empathy score was influenced by the degree of real
> similarity, and therefore constructed a more complicated dyadic
> index called a "refined empathy score" instead of isolating inde-
> pendent variables. This merely buried the artifacts deeper, as
> Hastorf and Bender ultimately discovered.

Cronbach suggested that investigators reduce their measures to these

components to clarify such problems.

Following Cronbach's recommendations, Cline (1958), using

colored sound films of interviews, conducted a study of the variables

related to accuracy in interpersonal perception, where the scoring

system used "was specifically chosen to make possible the more

analytical study of accuracy scores suggested by Cronbach." Cline

found that the ability to judge responses of others is a general but factorially complex process. He also found that Stereotype Accuracy accounted for a large portion of the generality: "Perhaps the most striking finding from the [Cline's] analysis is the importance of the stereotype accuracy component. The tendency in previous investigations has been however, to eliminate this component from accuracy scores in order to have 'purer' measures . . . . This seems to be an inappropriate emphasis . . . . It is really quite a surprising finding that having an accurate stereotype contributes so much to making accurate predictions about so diverse a group. It would therefore seem more appropriate to do research aimed at understanding Stereotype Accuracy than to eliminate it from consideration" (Cline, 1958).

PROBLEM

The measure of stereotype accuracy suggested by Cronbach and used by Cline is derived as a subscale from a film test that also provides measures of other components affecting accuracy. Because of the apparent importance of stereotype accuracy, the aim of the present study was to develop a separate test of this component, and to explore some of its possible correlates.

Speroff and Kerr (1951) made one of the earliest attempts to develop such a measure. Expressing dissatisfaction with earlier conceptions of empathy which neglected the "prediction of behavior" aspect, and focusing on the feelings and momentary experience of the subject, they developed a test designed to measure more objective and stable aspects of empathic ability. The Empathy Test (TET) consists of three series of items which the subject is asked to rank in popularity for various reference groups. Fourteen types of music are ranked by the subject in order of preference of "office workers"; 15 magazines in order of circulation rates; 10 annoying habits in order of their nuisance value for "persons over 40 years." Reliabilities by test-retest for this instrument range from .67 to .83.

The authors present the results of correlation studies supporting the validity of scores on their test: sociometric interpersonal desirability (r .34), supervisory aptitude as measured by How Supervise? (r's .62 and .32), campus offices held and number of semesters in

college (r's .15), breadth of smile of degree candidates receiving

their diplomas (r .64). It is the lone pencil and paper test in the

literature. But Thorndike (in Buros, 1959), says that ". . . the

relatively few studies [Bell, 1955; Siegel, 1954; and Van Zelst, 1953]

by persons not associated with the author have tended to yield pre-

dominately negative results. Unless the positive results reported

in the manual of the Empathy Test are verified in findings of other

workers, this test cannot be recommended as either a useful practical

device or a contribution to the description and understanding of the

individual. " It was the purpose of the present study, therefore, to

develop a test that would measure more general stereotypes than those

measured by Speroff and Kerr (1951), and that would be more valid.

METHOD

The manner in which each of the stereotype accuracy scales was developed and the subjects employed are first considered. This is followed by a discussion of the preliminary, experimental and final forms of that test. The concluding sections discuss other measures used in the study and correlated with the stereotype accuracy scales.

The College Man Scale

The items for this subscale were drawn from the Strong Vocational Interest Blank. E. K. Strong compared the interests of a large number of persons in regard to occupations, amusements, and activities. He tabulated the results in percentages for each activity, amusement or interest. Each item in this scale consisted of four interests. The correct answer for an item was that listed by Strong as having the highest preference. For example, for college men in general the percent who liked the following interests were:

Advertiser	53%
Politician	40%
Surgeon	46%
Aviator	65%
Poet	19%

Thus, several four-choice multiple-choice items can be generated from these five elements. Of the four choices included in any item,

that choice (on the basis of Strong's figures), which has the highest

per cent as being liked, is the correct answer.

The Woman Stereotype Scale

The scale for the Woman Stereotype was developed in the

same manner as was the scale for the College Man Stereotype. The

only difference being that figures for the average woman were used as

the criteria for correct answers.

The Sex Differences Scale

The scale for the sex stereotype again was the same as

for the previous two scales. The criterion used in selecting which

element of an item is the correct answer, however, differs. In this

subscale, the subject picks that choice which is most popular among

women and the least popular with men. That is, he must choose that

answer which shows the greatest sex difference. Here again Strong's

figures are used as a reference. This scale is best defined as

measuring Differential Stereotype.

The Age Stereotype Scale

The Age Stereotype subscale was constructed similarly

to the College Man scale and the average Woman scale except that

age levels were used as choices instead of interest activities. In this

scale only one activity or interest was given in each item. The subject's

task is to answer by giving the age group for which he thinks that activity is most popular. The possible age groups from which the subject chooses are 15, 25, or 55 year-olds. The age categories were also drawn from Strong.

The Subjects

All subjects in this study were students at Michigan State University. The various groups used as subjects were tested during their regular class periods. Table 1 gives a summary of the subjects in each group, and the data obtained from them. The total number of students in each group is given as well as the per cent males in each group.

Table 1

Subjects

Class	Term	N	% males	Data obtained
Industrial Psych.	Fall 1959	89	84	Prelim. form
Industrial Psych.	Winter 1959	100	85	Experimental form
Industrial Psych.	Spring 1959	77	84	Final form
Industrial Psych.	Spring 1959	77	84	Final form (retest)

Preliminary Form of College Man Scale

As a pilot to the main study, the first form was prepared, using 45 items. This form was administered to 80 students in the 1959 fall term Industrial Psychology course at Michigan State University. The scores obtained ranged from 10 to 39 with a mean of 24.95, a standard deviation of 5.88, and an internal consistency of r .66. The test was readministered to 70 of the original 89 students three weeks later and a test-retest correlation of .64 obtained. The items in this form were subjected to an item analysis, and the 30 best items were retained. These items constituted the College Man Stereotype subscale which was used in later forms in the study.

Experimental Forms

The preliminary form used in the pilot study was lengthened by adding the subscales described earlier (Woman Stereotype, Sex Stereotype, Age Stereotype). The Woman Stereotype subscale had 60 items as did the Sex Stereotype subscale. The 30 retained items from item analysis of the preliminary form constituted the College Man Stereotype subscale. Thus, the total number of items for the experimental form was 198. This form was administered to 140 students in the elementary course of Industrial Psychology at Michigan State University. Each scale was item analyzed separately to select the 30 most discriminating items. The highest and lowest scoring 27% of the students (for each scale separately) constituted the dichotomy

used as the basis for the item analysis. The 30 items which showed

the greatest discrimination between the high and the low scoring groups

were retained, and a new scoring key was developed for the full scale

of 120 items. The answer sheets were then rescored with the new key.

The 100 students who took the experimental form of the

test also took several other tests during the term as part of an ongoing

research project of which this thesis is a part.

Final Form

The 120 items (30 in each of four scales) selected by

application of item analysis to the experimental form, constitutes the

final form (see Appendix B). The order for the correct answers within

each subscale was randomized to control for response set, where a

subject may have a tendency to favor choosing a certain element

number of a series of items. In the four-choice items each choice is

correct in approximately one-fourth of the items, and in the three-

choice items each choice is correct one-third of the time. The average

time taken to complete the entire scale is from 30 to 40 minutes. The

scales may be machine or hand scored, using IBM scoring keys.

The final form was administered to an Industrial Psychology

class, spring term. Seventy-seven students took the test at the

beginning of the course and six weeks later.

Additional Measures

There were 13 additional measures available from the 100 winter term Industrial Psychology students. The relationships between these measures and the scales were investigated to explore further the nature of stereotype accuracy. The stereotype accuracy scores used in the analysis were those obtained from the rescoring of the best 120 items of the experimental form. The measures obtained are:

1.   Age.

2.   Sex.

3.   Grade in present Industrial Psychology course, consisting of the sum of the first two examination scores.

4.   American Council on Education Psychological Examination for College Freshmen. The total score served as a measure of the subject's general intelligence (Michigan State University, 1954-59).

5.   Trumbo Test of the Ability To Predict Behavior. This test contains six cases, each consists of a few sentences describing an actual person. The subject predicts the behavior of this person by answering 15 true-false statements about him. Actual behavior is used as a scoring criterion. Corrected odd-even reliability of this 120-item test was r .77 (Trumbo, 1955).

6.   Suppression-repression scale. An unpublished scale measuring the extent to which a person admits unpleasant or unfavorable thoughts and behavior. Items are similar to the MMPI L (lie) scale.

Validity and reliability for this scale have not been determined. This was one of two criterion measures used in this study.

7.   Cline Movie Prediction Test. This test is similar to the Trumbo Prediction Test, except that the subject sees a short sound color motion picture of a person being interviewed. The subject answers several types of items predicting the behavior of the person. The reliability of the test was .71, computed from the Spearman-Brown formula based on variances (Cline, 1958). Actual behavior is used as the scoring criterion. This test was used as a criterion measure for the present study.

8.   Scientific Values Scale. An unpublished scale which measures the extent to which a person believes in the value of science. Internal consistency for this scale was r .78, and the repeat reliability after six weeks was also r .78 (Hershey, 1958).

9.   Doré Leadership Questionnaire. This is a test designed to measure attitudes towards certain principles of leadership. It is a four scale test with 30 items in each scale, items being of the forced-choice type. Reliability was r .85 (Doré, 1960).

10.   Employee-Orientation. A subscale of the Doré leadership test which measures the extent to which the subject thinks a leader should consider the worker as an individual rather than as an instrument to get out production. Reliability: r .78.

11.   Delegation of Authority. Another subscale of the Doré test which measures the extent to which the subject thinks the leader

should set up general conditions for workers and let them figure out how the details of the work will be handled. Reliability: r .86.

12.   Differentiated Role.   The third of the four subscales of the Doré Leadership Questionnaire which measures the extent to which the respondent believes that the leader should or should not perform the same functions as the workers, plan the work, organize the work, or help train the workers. Reliability for this scale was r .75.

13.   Creates Teamwork.   The fourth of Doré's leadership test subscales.   This measures the extent to which the subject thinks a leader should work towards creating a spirit of teamwork within the work group, has pride in the group, and feels satisfaction with the accomplishments of the group.   Reliability of this subscale was r .68.

RESULTS

Below are summarized data dealing with the scales and their relationships to each other. Following this are the results having to do with their relationships to other measures. Then the results of the validity tests are given. Finally, norms for college students are reported.

Internal Consistency of the Subscales

The internal consistencies of the scales are reported below. These were calculated by the Kuder-Richardson formula #20, using data from the second administration of the 77 students in the spring term Industrial Psychology class.

College Man Stereotype	r .51
Woman Stereotype	r .16
Sex Stereotype	r .42
Age Stereotype	r .58
Total score	r .30

Stability of the Subscales

The test-retest reliabilities for the spring term class of Industrial Psychology students, with six weeks between administration, are given below. There were 77 students in this group.

College Man Stereotype $\underline{r}$. 60

Woman Stereotype $\underline{r}$. 44

Sex Stereotype $\underline{r}$. 46

Age Stereotype $\underline{r}$. 37

Total score $\underline{r}$. 56

Correlations between the Scales

Intercorrelations of the scales were computed for the winter term Industrial Psychology students (N = 100). The table below summarizes these findings.

Table 2

Intercorrelations of the Scales

	1	2	3	4
1. College Man Stereotype	-			
2. Woman Stereotype	. 11	-		
3. Sex Stereotype	-. 19	-. 08	-	
4. Age Stereotype	-. 01	. 00	. 28**	-

**Significant at . 01 level (critical level for . 01 = . 26)

Appendix E reports the complete table of intercorrelations for all measures obtained from the winter term Industrial Psychology students.

Factorial Structure

In an effort to determine the simple structure underlying the 18 measures obtained from the 100 winter term Industrial Psychology students, a principal axis factor analysis, using a rotational procedure suggested by Kiel and Wrigley (1960 APA Convention paper) was performed. Quartimax solutions were used. The procedure involves comparing solutions with increasing numbers of factors rotated until stability is reached. The factor analysis and the correlations between the 18 measures were computed on the MISTIC (Michigan State Iterative Computer). Five factors were identified, two of which concern the Stereotype Accuracy Test. Course grade, ACE scores, the Trumbo test, College Man Stereotype, and Woman Stereotype were the major variables identifying Factor I while Sex Stereotype, Age Stereotype and the total scores for the Stereotype Accuracy Test were factors defining Factor II. Table 3 reports the factor loadings for all variables on these two factors. The complete table of loadings on all variables for all factors can be found in Appendix C.

Correlations with Trumbo and Cline Tests

The Trumbo prediction test correlated .20 with the total score of the Stereotype Accuracy test. This was found to be significant at the .05 level. Subscale correlations with the Trumbo test were: College Man Stereotype .14, Woman Stereotype .08, Sex Stereotype .02, and Age Stereotype .13.

Table 3

Factor Loadings for All Eighteen Variables on the Two Factors

Defined by the Stereotype Accuracy Scales

Variable	Factor	
	I	II
College Man Stereotype	.55	.20
Woman Stereotype	.61	.20
Sex Stereotype	-.37	.58
Age Stereotype	-.09	.81
Stereotype Accuracy, Total score	.32	.89
Age	-.09	-.31
Sex	.12	-.04
Course Grade	.48	-.02
ACE	.62	.21
Trumbo	.33	.23
Suppression-Repression	.07	.00
Cline Movie	.04	.45
Scientific Values	-.24	.18
Differentiated Role	.25	-.14
Employee Oriented	-.01	.03
Delegation of Authority	-.08	-.01
Creates Teamwork	.01	.08
Leadership, Total score	.05	-.01

The total score of the Stereotype Accuracy Test correlated

.29 with the Cline Movie. Subscale correlations were found to be:

College Man Stereotype .02, Woman Stereotype .18, Sex Stereotype

.07, and Age Stereotype .32. Age Stereotype and the Cline Movie

correlated significantly beyond the .01 level, as did the total score

for the Stereotype Accuracy Test.

Correlations with Course Grades

A positive correlation was found between course grade in

the winter term Industrial Psychology group and the Woman Stereotype

subscale. Subscale correlations with course grade were: College Man

Stereotype .07, Woman Stereotype .20, Sex Stereotype .07, Age

Stereotype .02, and Total score .14.

Correlations with Doré Leadership Test

Certain subscales of the Doré Leadership Test correlated

significantly with some of the subscales of the Stereotype Accuracy Test.

They are reported in this section.

Differentiated Role. The Differentiated Role subscale of the

Doré Leadership scale with the College Man Stereotype was found to be

.20. This is significant at the .05 level.

Delegation of Authority and Teamwork. These two subscales

correlated positively with the Sex Stereotype subscale. Delegation of

Authority correlated .26; significant at the .01 level. Creates Teamwork

correlated .22; significant at the .05 level.

The table below summarizes all the findings of significant correlations between scores on the Stereotype Accuracy Test and other measures.

Table 4

Significant Correlations between the Stereotype
Scales and Other Variables

Related factors	CMS	WS	SS	AS	Total
Age			-.21*		
Sex	-.25*				
Course grade		.20*			
ACE		.27**			.28**
Cline Movie				.32**	.29**
Trumbo					.20*
Differentiated Role	.20*				
Delegation of Authority	-.24*		.26**		
Creates Teamwork			.22*		
Leadership Total			.21*		

$\underline{r}$ .05 = .20 $\underline{r}$ .01 = .26

Norms

Norms for the final form of the test were developed, and are
summarized in Table 5. The norms reported here were developed from
data gathered from the spring term Industrial Psychology class at
Michigan State University. This class may be considered a represen-
tative sample inasmuch as there was only one such section taught that
term, and there is reason to believe that this group was typical of the
groups taking this course other terms. (See Table 1 also.)

Table 5

Norms for Industrial Psychology Students

N = 77

Scale	First administration		Second administration	
	Mean	S. D.	Mean	S. D.
College Man Stereotype	16. 18	3. 60	16. 35	4. 26
Woman Stereotype	11. 70	2. 81	11. 57	2. 65
Sex Stereotype	13. 16	3. 24	12. 74	3. 24
Age Stereotype	18. 39	2. 72	17. 79	3. 65
Total	59. 42	5. 97	58. 43	9. 35

DISCUSSION

In spite of the weaknesses in measurement, some results are consistent with findings from other studies and lend support to the validity of the scales. Cline (1958) found that stereotype accuracy was a major determinant of scores on his test; the present study found a correlation of .29 between Cline scores and stereotype accuracy scores. Consistently, a correlation of .20 was found between Trumbo's test and stereotype accuracy scores. As in many tests of prediction ability, the present study reports a correlation of .28 between intelligence and stereotype accuracy. Finally, Cowden's (1955) finding that males are superior to females in ability to predict, is also supported by the correlation of -.25 between sex of the respondent and the College Man Stereotype.

The findings of the present study, reviewed below, lend little support to the still present possibility of finding a general, stable, and valid measure of stereotype accuracy. They do, however, suggest a new component of the problem as well as some further avenues to be explored.

Stereotype scales had discouragingly low reliabilities. Test-tetest reliabilities ranged from .37 to .60; the total score had a test-retest reliability of .56. Kuder-Richardson formula #20 internal consistency reliabilities ranged from .16 to .58.

Scores on the subscales, which were expected to be at least fairly closely related, were practically independent. Thus, as measured by these scales, stereotype accuracy appears to be a highly specific ability. Item analysis, by use of total scores, may lead to a test of greater internal consistency. The significance of scores on a test derived in this way is an open question.

Differential Stereotype Accuracy as a New Component of Prediction Ability

This subscale measures the extent to which the respondent can predict interests which show the greatest sex difference in terms of preference. Inclusion of the scale has created theoretical and practical problems from the beginning of the study. The subscale was included primarily because of the convenient fact that Strong, in tabulating the interests of people, reported those interests which differentiated men and women most (Strong, 1943). The compatibility of this subscale with Cronbach's definition of Stereotype Accuracy, therefore, was not questioned until after most of the data had been gathered and analyzed. A routine re-check of some of the literature on Stereotype Accuracy revealed that the Sex Stereotype subscale was, in fact, not a measure of Stereotype Accuracy. A closer examination of Cronbach's analysis of the components of empathy was carried out in order to determine what actually had been measured by this subscale. Here is where an important finding was uncovered.

Cronbach (1955) breaks empathy down into four components. He used data from a study of a group of students at Cornell. His design was two dimentional. None of the definitions for the components of empathy in his analysis were suitable as models for what was measured by the Sex Stereotype scale. This was simply because, if the measure derived by the Sex Stereotype were applied to the Cronbach design, that design would need to be extended by another dimension. This is no criticism of Cronbach's design; he was not interested in any information that could be gained by extending his design. Using Cronbach's analysis as a paradigm, the Sex Stereotype subscale can be called Differential Stereotype and described as follows:

<u>Differential Stereotype</u> (DS). This is a new concept that can be considered a component of empathy. It is measured in this study by the Sex Stereotype subscale. The Sex Stereotype subscale is not a measure of Stereotype Accuracy, it is a measure of <u>Differential Stereotype</u>. It reflects the ability of the judge to predict the differences on any item (averaged over items) between groups of persons. This definition differs from the Cronbach definition of Differential Accuracy in that the Cronbach definition is concerned with differences between <u>individuals</u>. Here the concern is with differences between <u>groups</u>. This difference, averaged over items, is the factor which the judge must predict accurately. We may write:

$$DS_{.i}^2 = \sigma_{y''_{.ij}}^2 + \sigma_{x''_{.i}}^2 - 2\sigma_{y''_{.ij}} \sigma_{x''_{.i}} r_{y''_{.ij} x''_{.i}} \qquad [3]$$

The variance is computed over items. $\sigma_{x"_{.i}}^{2}$ indicates the scatter of the actual means. This score depends on J's knowledge of the differences in popularity or relative frequency of responses for any item between groups. $y"_{.ij}$ is J's description of the differences between two groups on item i. $\sigma_{y"_{.ij}}$ expresses J's report of how much one group of Os will differ from another group of Os in responding to an item. It is the assumed similarity between groups. $r_{y"_{.ij}x"_{.i}} = DS_r$ and reflects J's ability to judge which group of Os have the highest scores on the item.

The concept may be clarified by considering the stereotypes popularly attributed to Negroes and Jews. The same characteristics are not attributed to both groups. Rather, differences are recognized in the characteristics claimed of both groups, some of these are accurate and others are not. Thus, it is conceivable that differences between the responses of these two groups on any one item can be predicted. This in fact is what was done in the case of SS, where males and females were the dichotomy in question.

From our knowledge of analysis of variance, we know that the component variances add up to one. Hence in the extended dimension analysis of variance paradigm offered here, Differential Stereotype is a component of empathy in the same sense as are the four analyzed by Cronbach. In the present study no analysis of variance was actually calculated.

The implications of this finding are several.

If the analysis offered in this study regarding Sex Stereotype is correct, this would suggest that there are at least five components of empathy, and perhaps more, depending on the complexity of the analysis calculated.

The possibilities for combinations of groups whose differences would be predicted are almost without limit. For example, the differences between labor and management, Democrats and Republicans, easterners and westerners, and men and women on a variety of issues, as well as men and women in general.

SUMMARY AND CONCLUSIONS

The purpose of this study was to develop an objective and economical test that would measure Stereotype Accuracy, the ability to predict the behavior of social groups.

Four 30-item scales were designed to measure the accuracy of four stereotypes: College Man Stereotype, Woman Stereotype, Sex Stereotype, and Age Stereotype. These were developed from four preliminary subscales, using item analysis of the longer form. The items in the test were multiple-choice items, with elements being drawn from the Strong Vocational Interest Blank. The correct answer for each item was that listed by Strong as having the highest preference by persons of a given category (stereotype). Repeat reliabilities of the stereotype scales and total scores ranged from .37 to .60. The internal consistency, as measured by the Kuder-Richardson formula #20, ranged from .16 to .58.

Thirteen additional measures were obtained from one group of 100 students to explore the relationships of the measures to the scores on the several scales. A total of 266 college students served as subjects in the various phases of the study.

Evidence was found to support the following conclusions. The accuracy of a person's stereotype is related to his accuracy in predicting the behavior of individuals (r of .20 in one test, and r of .29 in another). The accuracy of a person's stereotype has little relation

to his knowledge of psychology (r of .14 with course grade in psychology).
The accuracy of a person's stereotype is related to his attitudes towards
methods of successful leadership (correlations ranged from -.24 to .26).

As measured by the present instrument, stereotype accuracy
appeared to be a highly specific ability, for the correlations between the
subscales ranged from -.19 to .28. Whether this is an artifact of the
measuring instrument and whether instruments with greater generality
can be developed, are questions that can be resolved only by further
experimental inquiry.

Several suggestions for future research emerge from this
study. First among them would be a program aimed at improving
reliabilities and validities of the instrument. For example, a correlational
study with the Speroff and Kerr Empathy Test would be of value.

Once the present test has been developed to the point that
it would be of service in applied as well as theoretical problems, a
series of tests, each measuring a separate component of empathy would
be highly desirable. For example a test measuring Differential
Stereotype, one measuring Differential Accuracy, and so on, would be
of value in re-investigating some of the earlier studies on empathy.
Perhaps in this way a better understanding of the process of empathy
can be attained.

REFERENCES

Allport, G. W. Personality, Ch. XIX, p. 532, New York: Holt & Co.,

1937.

Anikeef, Alexis M. Sociometric empathy. J. Social Psychology, 1957,

45, 283-287.

Baxter, Nathan. A study of the relationships between self-awareness

and social sensitivity. Dissertation Abstracts, 1952,

12, 719-720.

Baumgarten, F. Sympathy, understanding and empathy. Psychological

Abstracts, 1931, No. 241.

Bell, Graham B., & Stolper, R. An attempt at validation of the

empathy test. J. Applied Psychology, 1955, 39, 442-443.

Bell, G. B., & Hall, H. E. Jr. The relationship between leadership

and empathy. J. Abnormal & Social Psychology, 1954,

49, 156-157.

Bender, I. E. & Hastorff, A. H. On measuring generalized empathic

ability (social sensitivity). J. Abnormal & Social Psychology,

1953, 48, 503-506.

Buros, O. K. (Ed.). Fifth mental measurements yearbook. Highland

Park, New Jersey: Gryphon Press.

Cline, V. B., & Richards, J. M. Variables related to accuracy in

interpersonal perception. First Annual Report, November,

1958, Contract No. NR 171-146, Group Psychology Branch,

Office of Naval Research, University of Utah.

Cowden, R. C. Empathy or projection? J. Clinical Psychology, 1955, 11, 188-190.

Cronbach, L. J. Processes affecting scores on understanding of others" and "assumed similarity." Psychological Bulletin, 1955 52, 177-193.

Crow, W. J., and Hammond, K. R. The generality of accuracy and response sets in interpersonal perception. J. of Abnormal and Social Psychology, 1957, 54, 384-390.

Crow, W. J. In defense of stereotypes, paper read at 1960 convention of APA, Chicago, Ill.

Dailey, C. Unpublished Ph. D. dissertation, U. of M., 1950.

Doré, R. L. The development and validation of forced choice scales measuring attitudes toward leadership methods. Unpublished master's thesis, M. S. U., 1960.

Dymond, R. A preliminary investigation of the relation of insight and empathy. J. Consult. Psychol., 1938, 12, 228-233.

Dymond, R. Personality and empathy. J. Consult. Psychol., 1950. 14, 343-350.

Escalona, S. Feeding disturbances in very young children. Amer. J. Ortho., 1945, 15, 76-80.

English, H. B. & English, A. C. A comprehensive dictionary of psychological and psychoanalytical terms. New York: Longmans, Green, 1958.

File, Q. W. The measurement of supervisory quality in industry.

J. Applied Psychol., 1945, 29, 323-337.

Fleiss, R. The metapsychology of the analyst. Psychiatric Quart.,

1942, 211-227.

Foa, U. G. The contiguity principle in the structure of interpersonal

relationships. Human Relations, 1958, 11, 229-238.

Foa, U. G. Some correlates of the workers with the foreman. J.

Applied Psychol., 1960, 44, 6-11.

Force, R. C., Empathy tests for selecting personnel to deal with

people in trouble. J. Soc. Ther., 1956, 2, 46-50.

Fromm-Reichmann, F. Principles of intensive psychotherapy. Chicago:

University of Chicago Press, 1950.

Gage, N. L. Explorations in the understanding of others. Educ.

Psychol. Measmt., 1953, 13, 14-26.

Gage, N. L., Leavitt, G. S. and Stone, G. C. The intermediary key in

the analysis of interpersonal perception. Psychol. Bull.,

1956, 62, 411-461.

Guilford, J. P. Fundamental statistics in psychology and education.

New York: McGraw-Hill, 1956.

Halpern, H. M. Some factors involved in empathy. Dissertation

Abstracts, 1954, 14, 1957.

Hastroff, A. H. and Bender, I. E. A caution respecting the measure-

ment of empathic ability. J. Abnorm. Soc. Psychol.,

1952, 47, 574-576.

Hastorff, A. H. & Bender, I. E. On measuring empathic ability (Social sensitivity). J. Abnorm. & Soc. Psychol., 1953, 48, 503-506.

Heidbreder, E. Seven psychologies. New York: Appleton-Century, 1933, p. 101.

Hershey, G. L. College grades in relation to inventory measures of personality. Unpublished master's thesis. Michigan State University, 1958.

Hutt, M. & Miller, D. Social values and personality development, J. Social Issues, 1949, 5, No. 4.

Jarrard, L. E. Empathy: The concept and in-trial applications. Personnel Psychology, 1956, 9, 157-158.

Johnson, R. J. Relationship of employee morale to ability to predict responses. J. Applied Psychol., 1954; 38, 32-323.

Kerr, W. A. & Speroff, B. J. The empathy test. Chicago, Ill.: Psychometric Affiliates, 1951.

Lipps, T. Das Wissen von Frenden Ichen. Psychologische Untersuchungen, 1907, v. I, 694-722.

Luft, J. Differences in prediction based on hearing versus reading verbatim clinical interviews. J. Consult. Psychol., 1951, 15, 115-119.

Luszki, M. Empathic ability and social perception. Ph. D. dissertation. Dissertation Abstracts, University of Michigan, 1951.

McClelland, W. A. Preliminary test of role playing ability. J. Consult. Psychol., 1951, 15, 102-108.

McMurry, R. N. Empathy: Management's greatest need. Advance Management, 1953, 18(7); 6-11; 34.

Miller, F. G. & Remmers, H. H. Studies in industrial empathy, II. Management's attitudes toward industrial supervision and their estimates of labor attitudes. Personnel Psych., 1950, 3, 33-40.

Murstein, B. I. Some comments on the measurement of projection and empathy. J. Consult. Psych., 1957, 21, 81-82.

Nagle, B. F. Productivity, employee attitude and supervisor sensitivity. Personnel Psychol., 1954, 7, 219-233.

Norman, R. D. & Leiding, W. C. The relationships and measures of individual and mass empathy. J. Consult. Psychol., 1956, 20, 79-82.

Patton, W. M. Jr. Studies in industrial empathy III: A study of supervisory empathy in the textile industry. J. Applied Psych., 1954, 38, 285-288.

Porter, L. W. & Kaufman, R. A. Relationships between a top-middle management self-description scale and behavior in a group situation. J. Applied Psych., 1959, 43: 345-348.

Remmers, H. H. A quantitative index of social-psychological empathy. J. Orthopsychiat., 1950, 20, 16-165.

Remmers, L. J. & Remmers, H. H. Studies in industrial empathy I: Labor attitudes toward industrial supervision and their estimate of management's attitudes. Personnel Psych., 1949, 2, 427-436.

Rogers, C. S. Client Centered Therapy. Boston: Houghton-Mifflin, 1950, p. 29.

Reusch, J. Experiments in psychotherapy, I Theoretical considerations. J. Psychiat., 1948, 25, 137-169.

Reusch, J. & Prestwood, A. Anxiety, its initiation, communication and interpersonal management, Arch. Neurol. & Psychiat., 1948, 62, 527-550.

Siegel, A. I. An experimental evaluation of the sensitivity of the empathy test. J. Applied Psych., 1954, 38, 222-223.

Speroff, B. J. Relationship between empathic ability and supervisory knowledge. J. Personnel Admin. in Industrial Relations, 1954, 1, 195-197.

Speroff, B. J. Empathic ability and accident rate among steel workers. Personnel Psychol., 1953, 6, 297-300.

Speroff, B. J. Empathy and role reversal as factors in industrial harmony. J. Soc. Psychol., 1953, 37, 117-120.

Stagner, R. Psychological aspects of industrial conflict. Personnel Psychol., 1948, 1, 131-143.

Stone, G. C., Gage, N. L. & Leavitt, G. S. Two kinds of accuracy in predicting another's responses. J. Soc. Psychol., 1957, 45, 245-254.

Strong, E. K.   Vocational interests of men and women.   Stanford:

Stanford University Press, 1943.

Sullivan, H. S.   Conceptions of modern psychiatry.   Washington, D. C.:

W. A. White Foundation, 1940, p. 6.

Taft, R.   The ability to judge people.   Psych. Bulletin, 1955, 52, 1-22.

Taguiri, R. & Petrullo, L.   Person perception and interpersonal

behavior.   Stanford: Stanford Univ. Press, 1958.

Thorndike, R. L.   Personnel selection.   New York: Wiley, 1949.

Trumbo, D. A.   The development and analysis of a test of the ability to

predict behavior.   Unpublished master's thesis.   Michigan

State University, 1955.

Van Zelst, R. H.   Empathy test scores of union leaders.   J. Applied

Psych., 1953, 36, 293-295.

Van Zelst, R. H.   Validation evidence on the empathy test.   Educ.

Meas., 1953, 13, 474-477.

Vernon, G. M. & Stewart, R. L.   Empathy in the dating situation.

Amer. Soc. Rev., 1957, 22(1), 48-52.

Wrigley, C. F., & Neuhaus, J.   The quartimax method: An approach

to the orthogonal simple structure.   Brit. J. Stat. Psychol.

Nov., 1954.

APPENDIX A

H. C. Smith
January, 1960

STEREOTYPE ACCURACY

Much of our ability to judge a specific person depends on our understanding of the general type to which the person belongs: Our understanding of a particular college man depends on our understanding of college men in general; our understanding of a particular woman depends on our understanding of women in general; our understanding of the differences between a particular man and woman depends on our understanding of the differences between men and women in general; our understanding of the differences between a particular young man and particular old man depends on our understanding of the differences between young and old men in general; etc.. The four tests below measure your general understanding of the interests of women and college men, of the different interests of men and women, and of the different interests of young and old people.

College Man Stereotype

Undergraduate men in a Midwestern university were asked whether they liked or disliked each of a large number of different occupations, school subjects, amusements, and kinds of people. Mark on the separate answer sheet which one of the four in the following groups you think was liked by the most students.

1. (1) artist (2) army officer (3) author of a novel (4) pharmacist.
2. (1) surgeon (2) mechanical engineer (3) statistician (4) interpreter.
3. (1) astronomer (2) athletic director (3) sculptor (4) toolmaker.
4. (1) watchmaker (2) rancher (3) landscape gardener (4) mining superintendent.
5. (1) undertaker (2) cashier in bank (3) music teacher (4) dentist.
6. (1) magazine writer (2) chemist (3) retailer (4) secret service man.
7. (1) ship officer (2) carpenter (3) laboratory technician (4) railway conductor.
8. (1) printer (2) scientific research worker (3) judge (4) orchestra conductor.
9. (1) laboratory technician (2) politician (3) private secretary (4) toolmaker.
10. (1) photo engraver (2) jeweler (3) ship officer (4) explorer.
11. (1) poet (2) printer (3) scientific research worker (4) governor of a state.
12. (1) secret service man (2) farmer (3) auto racer (4) inventor.
13. (1) army officer (2) electrical engineer (3) bookkeeper (4) aviator.
14. (1) history (2) calculus (3) algebra (4) chemistry.
15. (1) nature study (2) physics (3) economics (4) military drill.
16. (1) geology (2) zoology (3) geometry (4) philosophy.
17. (1) English composition (2) mechanical drawing (3) mathematics (4) psychology
18. (1) art (2) geography (3) physics (4) zoology.
19. (1) military drill (2) shop work (3) public speaking (4) calculus.
20. (1) civics (2) geology (3) dramatics (4) languages, ancient.
21. (1) "New Republic" (2) "System" (3) "Time" (4) "Popular Mechanics"
22. (1) musical comedy (2) symphony concerts (3) vaudeville (4) fortune tellers.
23. (1) solving mechanical puzzles (2) full dress affairs (3) auctions
    (4) performing sleight-of-hand tricks.
24. (1) repairing electrical wiring (2) cabinet making (3) interviewing clients
    (4) climbing along edge of precipice.
25. (1) decorating a room with flowers (2) contributing to charity (3) adjusting
    a carburetor (4) adjusting difficulties of others.

26. (1) continually changing activities (2) displaying merchandise in a store (3) expressing judgments publicly regardless of criticism (4) looking at a collection of antigue furniture.
27. (1) repairing a clock (2) looking at a collection of rare laces (3) pursuing bandits in sheriff's posse (4) entertaining others.
28. (1) foreigners (2) fashionably dressed people (3) independents in politics (4) energetic people.
29. (1) conservative people (2) optimists (3) people with gold teeth (4) talkative people.
30. (1) people who chew gum (2) people who talk very slowly (3) people who are natural leaders (4) thrifty people.

THE WOMAN STEREOTYPE

E. K. Strong[3] asked several thousand women whether they liked the activities involved in a variety of occupations and amusements, and certain types of people. He then calculated the percent of women who liked each. Mark on the separate answer sheet the one of the four interests in each of the groups below that you think was liked by the most women.

31. (1) Architect (2) advertiser (3) accountant (4) movie actress.
32. (1) Stage actress (2) artist (3) author of children's books (4) aviatrix.
33. (1) artist's model (2) athletic director (3) author of novel (4) bank teller.
34. (1) buyer of merchandise (2) biologist (3) bacteriologist (4) bookkeeper.
35. (1) costume designer (2) cook (3) confectioner (4) cashier.
36. (1) dietitian (2) Dean of Women (3) dramatist (4) florist.
37. (1) cartoonist (2) caterer (3) chemist (4) foreign correspondent.
38. (1) college professor (2) employment manager (3) editor (4) interior decorator
39. (1) interpreter (2) inventor (3) landscape gardner (4) judge.
40. (1) magazine writer (2) librarian (3) laboratory technician (4) illustrator.
41. (1) opera singer (2) music composer (3) musician (4) hostess.
42. (1) physician (2) minister (3) nurse (4) naturalist.
43. (1) playground director (2) poet (3) office manager (4) milliner.
44. (1) professional dancer (2) private secretary (3) psychiatrist (4) publisher.
45. (1) psychologist (2) social worker (3) surgeon (4) eductional director.
46. (1) musician (2) wife (3) interior decorator (4) psychologist.
47. (1) vocational counselor (2) tea room proprietor (3) sculptress (4) poet.
48. (1) dancing (2) camping (3) playing a musical instrument (4) animal zoos.
49. (1) riding horses (2) golf (3) swimming (4) bridge.
50. (1) poetry (2) conventions (3) driving an automobile (4) afternoon teas.
51. (1) picnics (2) tennis (3) formal affairs (4) fortune tellers.
52. (1) dancing (2) swimming (3) plays (4) picnics.
53. (1) art galleries (2) amusement parks (3) conventions (4) attending lectures.
54. (1) poker (2) conventions (3) animal zoos (4) art galleries.
55. (1) museums (2) golf (3) detective stories (4) fortune tellers.
56. (1) riding horses (2) musical comedy (3) bridge (4) observing birds.
57. (1) symphony concerts (2) bridge (3) poker (4) afternoon teas.
58. (1) fortune tellers (2) animal zoos (3) attending lecturers (4) movies.
59. (1) "House and Garden" magazine (2) romantic stories (3) movie magazines (4) poetry.
60. (1) "National Geographic Magazine" (2) "Vanity Fair" (3) "True Story" (4) poetry.
61. (1) "Reader's Digest" (2) "New Republic" (3) "House and Garden" (4) "True Story"
62. (1) Energetic people (2) conservative people (3) people who assume leadership (4) independents in politics.
63. (1) people who take life seriously (2) foreigners (3) optimists (4) athletic women.

64. (1) foreigners (2) religious people (3) methodical people (4) cautious people.
65. (1) emotional people (2) people who are natural leaders (3) self-conscious people (4) fashionably dressed people.
66. (1) people who are unconventional (2) men who drink (3) people who have made fortunes in business (4) thrifty people.
67. (1) people who take chances on situations of doubtful outcome (2) athletic women (3) thrifty people (4) people who have done you favors.
68. (1) religious people (2) people who are unconventional (3) people who take life seriously (4) sick people.
69. (1)negroes (2) sick people (3) methodical people (4) fashionably dressed people.
70. (1) very old people (2) absent-minded people (3) cautious people (4) people who tell you their troubles.
71. (1) men who drink (2) conservative people (3) nervous people (4) people with physical disabilities.
72. (1) calculus (2) algebra (3) physics (4) shorthand.
73. (1) art (2) bible study (3) bookkeeping (4) journalism.
74. (1) chemistry (2) civics (3) English composition (4) political science.
75. (1) geography (2) geology (3) geometry (4) domestic science.
76. (1) sociology (2) history (3) public speaking (4) physical training.
77. (1) modern languages (2) dramatics (3) ancient languages (4) physics.
78. (1) physiology (2) literature (3) art (4) philosophy.
79. (1) music (2) botany (3) economics (4) sociology.
80. (1) physics (2) geology (3) mechanical drawing (4) nature study.
81. (1) philosophy (2) penmanship (3) public speaking (4) typewriting.
82. (1) statistics (2) political science (3) physical training (4) shorthand.
83. (1) chemistry (2) physiology (3) zoology (4) chemistry.
84. (1) arithmetic (2) psychology (3) economics (4) English composition.
85. (1) spelling (2) journalism (3) teacher training (4) typewriting.
86. (1) decorating a room with flowers (2) attending church (3) preparing dinner for guests (4) giving "first-aid" assistance.
87. (1) reading editorial columns (2) buying at an auction sale (3) saving money (4) taking responsibility.
88. (1) raising money for a charity (2) entertaining others (3) meeting new situations (4) contributing to charities.
89. (1) trying new cooking recipes (2) looking at shop windows (3) meeting and directing people (4) writing a personal letter.
90. (1) regular hours of work (2) writing reports (3) making a speech (4) cooking.

Sex Stereotypes

E. K. Strong compared the interests of a large number of women and men. About the same number of women and men expressed a liking for most occupations, amusements, and activities. But for some interests more women expressed a liking than men. Mark on the separate answer sheet the interest in each of the groups below that showed the greatest sex difference. Mark, in other words, the one interest that you think the most women liked and the fewest men.

91. (1) artist (2) accountant (3) advertiser (4) architect.
92. (1) bacteriologist (2) athletic director (3) author of novel (4) bank teller.
93. (1) biologist (2) bookkeeper (3) cartoonist (4) cashier.
94. (1) buyer of merchandise (2) cook (3) chemist (4) confectioner.
95. (1) costume designer (2) Dean of Women (3) dietitian (4) florist.
96. (1) employment manager (2) factory manager (3) factory worker (4) farmer
97. (1) educational director (2) foreign correspondent (3) editor (4) hotel manager.
98. (1) interior decorator (2) housekeeper (3) illustrator (4) dress maker.
99. (1) lawyer (2) judge (3) interpreter (4) laboratory technician.
100. (1) missionary (2) minister (3) milliner (4) landscape gardener.

101. (1) musician (2) opera singer (3) music composer (4) museum director.
102. (1) playground director (2) physician (3) pharmacist (4) office manager.
103. (1) proof reader (2) poet (3) radio singer (4) professional dancer.
104. (1) social worker (2) probation officer (3) psychologist (4) psychiatrist.
105. (1) reporter (2) registrar (3) scenario writer (4) private secretary.
106. (1) social worker (2) telephone operator (3) art teacher (4) statistician.
107. (1) vocational counselor (2) high school teacher (3) stenographer
     (4) art teacher.
108. (1) typist (2) stenographer (3) telephone operator (4) commercial teacher.
109. (1) taking long walks (2) dancing (3) swimming (4) camping.
110. (1) golf (2) tennis (3) riding horses (4) driving an automobile.
111. (1) bridge (2) poker (3) playing a musical instrument (4) afternoon teas.
112. (1) animal zoos (2) amusement parks (3) picnics (4) conventions.
113. (1) plays (2) movies (3) fortune tellers (4) formal affairs.
114. (1) museums (2) attending lectures (3) art galleries (4) observing birds.
115. (1) musical comedy (2) romantic stories (3) detective stories (4) movie
     magazines.
116. (1) symphony concerts (2) plays (3) movies (4) attending lectures.
117. (1) movie magazines (2) detective stories (3) romantic stories (4) poetry.
118. (1) "Ladies Home Journal" (2) "National Geographic Magazine "
     (3) "Atlantic Monthly" (4) "House and Garden" magazine.
119. (1) operating machinery (2) giving "first-aid" assistance (3) attending
     church (4) making a speech.
120. (1) raising flowers and vegetables (2) being the first to wear the very latest
     fashions (3) repairing electric wiring (4) interviewing clients.
121. (1) arguments (2) making a speech (3) teaching adults (4) decorating a
     room with flowers.
122. (1) organizing a play (2) preparing dinner for guests (3) meeting and
     directing people (4) taking responsibility.
123. (1) teaching children (2) teaching adults (3) discussions of economic affairs
     (4) discussions of politics.
124. (1) writing reports (2) entertaining others (3) acting as yell-leader
     (4) adjusting difficulties of others.
125. (1) writing personal letters (2) buying at an auction sale (3) trying new
     cooking recipes (4) doing research work.
126. (1) saving money (2) regular hours of work (3) looking at shop windows
     (4) preparing dinner for guests.
127. (1) contributing to charities (2) saving money (3) taking responsibility
     (4) meeting new situations.
128. (1) expressing judgments publicly, regardless of criticism (2) raising
     money for charity (3) being head of a civic improvement program
     (4) discussions of politics.
129. (1) looking at a collection of rare laces (2) organizing a play (3) enter-
     taining others (4) contributing to charities.
130. (1) progressive people (2) conservative people (3) people who borrow things
     (4) energetic people.
131. (1) absent-minded people (2) very self-confident people (3) pessimists
     (4) people who are natural leaders.
132. (1) people who assume leadership (2) very intellectual people (3) emotional
     people (4) thrifty people.
133. (1) people who have done you favors (2) people who take life seriously
     (3) witty people (4) negroes.
134. (1) cautious people (2) sick people (3) people with physical disabilities
     (4) people who always agree with you.
135. (1) people who tell you their troubles (2) people who talk very loudly
     (3) people who talk about themselves (4) methodical people.

78. (1) contributing to charities (2) organizing a play (3) entertaining others (4) looking at a collection of rare laces.
79. (1) People who assume leadership (2) very intellectual people (3) emotional people (4) thrifty people.
80. (1) people who tell you their troubles (2) people who talk very loudly (3) people who talk about themselves.
81. (1) independents in politics (2) carelessly dressed people (3) "Mannish" women (4) fashionably dressed people.
82. (1) nervous people (2) people who chew gum (3) men who drink (4) women who smoke.
83. (1) athletic women (2) women cleverer than you are (3) cautious people (4) optimists.
84. (1) energetic people (2) religious people (3) very intellectual people (4) absent-minded people.
85. (1) thrifty people (2) self-conscious people (3) people who are natural leaders (4) foreigners.
86. (1) displaying merchandise in a store (2) being left to yourself (3) continually changing activities (4) saving money.
87. (1) making a speech (2) reading editorial columns (3) opening a conversation with a stranger (4) taking responsibility.
88. (1) president of a society (2) secretary of a society (3) treasurer of a society (4) member of a society.
89. (1) Chairman, Entertainment Committee (2) Chairman, Educational Committee (3) Chairman, Membership Committee (4) Chairman, Arrangement Committee.
90. (1) draftsman (2) factory worker (3) chemist (4) Buyer of merchandis

Age Stereotypes

E. K. Strong studied the changing interests of men from youth to old age. On the whole, he found that interests changed relatively little with age. He did find, however, that 15-year olds, 25-year olds, and 55-year olds had some distinctive interests. For each of the interests below mark:

    (1) If you think it was liked most by 15-year olds.
    (2) If you think it was liked most by 25-year olds.
    (3) If you think it was liked most by 55-year olds.

91. Civil engineer
92. Auto racer
93. Machinist
94. Auctioneer
95. Poet
96. Social worker
97. Statistician
98. Politician
99. Florist
100. Auto repairman
101. Life insurance salesman
102. Hunting
103. Thomas A. Edison
104. Popular Mechanics
105. Progressive people
106. Saving money
107. Energetic people
108. People who are natural leaders

166. Hunting
167. Athletic men
168. Fishing
169. Thomas A. Edison
170. Popular Mechanics
171. Progressive people
172. Operating new machine
173. Saving money
174. Energetic people
175. People who are natural leaders
176. Handling horses
177. Travel movies
178. Museums
179. Educational movies
180. Musical comedy
181. Sporting pages
182. Discussing my ideals with others
183. Thrifty people
184. People who talk very loudly
185. Decorating a room with flowers
186. Teaching adults
187. Economics
188. President of a society or club
189. Chairman, entertainment committee
190. Chairman, publicity committee
191. Looking at a collection of rare laces
192. Vaudeville
193. New Republic
194. Men who chew tobacco
195. Electrical engineer
196. Floor walker
197. Specialty salesman
198. Life insurance salesman

1-18-60   sl

109. Handling horses
110. Travel movies
111. Museums
112. Musical comedy
113. Discussing my ideals with others
114. Thrifty people
115. People who talk very loudly
116. Decorating a room with flowers
117. Economics
118. President of a society or club
119. Vaudeville
120. Floorwalker

2/16/60   æz

APPENDIX C

Rotated Factor Loadings for SA and Other Measures

Variable	I	II	III	IV	V
Age	-.036	-.086	-.314	-.454	-.240
Sex	-.136	.124	-.045	.803	-.148
Course Grade	-.434	.483	-.024	.066	.187
ACE	-.122	.615	.212	.108	.306
Trumbo Prediction	-.330	.333	.229	-.089	.308
S-I	.024	.074	.002	-.741	-.080
Cline Movie	-.015	.040	.453	.133	-.474
Sci. Val.	-.120	-.236	.180	-.092	.671
Differentiated Role	-.352	.250	-.138	-.247	-.490
Employee Oriented	-.870	-.010	.027	-.008	.040
Delegation of Authority	-.824	-.076	-.013	.206	.091
Creates Teamwork	-.791	.008	.082	-.024	.018
Total (Doré Test)	-.984	.053	-.010	-.004	-.091
College Man Stereotype	.112	.549	.199	-.458	-.072
Woman Stereotype	.010	.607	.204	.220	.004
Sex Stereotype	-.279	-.373	.576	.240	.065
Age Stereotype	.055	-.095	.813	-.049	-.117
Total (Stereotype Accuracy)	-.037	.319	.892	-.064	-.069

APPENDIX D

Notations

Notations in the text follow those of Cronbach, except for notation number 8, which is an adaptation from Cronbach's and makes up the formula used to identify the new concept introduced in this paper.

1. x_{oi} Self-description of Other $\underline{o}$ on item $\underline{i}$

2   y_{oij} Judge $\underline{j's}$ description of $\underline{o}$ on $\underline{i}$

3. $\bar{x}_o$ Average of x_{oi} over all items

4. $\bar{x}_{.i}$ Average over Others

5. $\bar{x}_{..}$ Grand mean over Others and items. $\underline{y}$ is defined similarly.

6. $x_{oi'} = x_{oi} - \bar{x}_{o.} - \bar{x}_{.i} + \bar{x}_{..}$ the score x_{oi} transformed as a deviation from both item mean and Other mean; y_{oij} similarly.

7. $(y_{oij} - x_{oi})^2$ is error in prediction

8. $y_{.ij''}$ Judge $\underline{j's}$ description of the differences between two groups on item $\underline{i}$.

APPENDIX E

Complete Table of Intercorrelations on All Measures for Winter Term Psychology Students.   N = 100

	Age	Sex	Grades	ACE	Trumbo	S-R	Cline	Sci. Val.	Diff. Role	Empl. Orient.	Deleg. Authority	Creates Teamw'k	Doré	Col. Man Stereo.	Woman Stereo.	Sex Stereo.	Age Stereo.	Stereo. Acc. Tl.
1.	-																	
2.	-24	-																
3.	04	16	-															
4.	-23	13	23	-														
5.	-12	-04	30	26	-													
6.	21	-45	-04	04	00	-												
7.	-06	03	-06	02	08	-10	-											
8.	-03	-12	11	05	05	-11	-07	-										
9.	15	00	27	04	00	15	04	-13	-									
10.	-05	09	23	12	33	-02	04	06	13	-								
11.	-08	19	33	11	25	-14	01	10	12	69	-							
12.	01	10	25	12	17	01	02	13	16	68	46	-						
13.	03	14	38	13	27	-02	04	06	41	87	82	78	-					
14.	01	-25	07	17	14	19	02	-07	20	-02	-24	-04	-05	-				
15.	-06	12	20	27	08	-06	18	-03	-08	00	02	12	-04	11	-			
16.	-21	16	07	00	02	-19	07	18	-03	17	26	22	21	-19	-08	-		
17.	-14	00	-02	13	16	08	32	05	-07	-05	-04	-01	-05	-01	00	28	-	
18.	-19	01	14	28	20	03	29	05	02	04	-02	12	06	47	44	47	66	-

*Critical value for .05 level is .20

**Critical value for .01 level is .26

Assume a decimal point in front of every digit pair