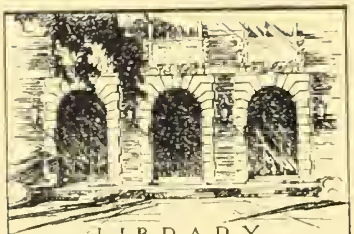

GROUP EFFECTIVENESS RESEARCH LABORATORY
Department of Psychology
University of Illinois
Urbana, Illinois

PROCESSES AFFECTING "UNDERSTANDING OF OTHERS" AND "ASSUMED SIMILARITY"

LEE J. CRONBACH
College of Education, University of Illinois

TECHNICAL REPORT NO. 10

Study performed under Contract N6-ori-07135
with the Office of Naval Research

PROJECT ON
SOCIAL PERCEPTION AND GROUP EFFECTIVENESS

AUGUST, 1954

PROCESSES AFFECTING SCORES ON "UNDERSTANDING OF OTHERS" AND "ASSUMED SIMILARITY"*

Lee J. Cronbach

College of Education, University of Illinois

How one person judges another is important both for its theoretical implications and for its practical significance in leadership, clinical assessment, and teaching skill. Recent studies of "social perception", as this area is termed, have been chiefly concerned with differences among perceivers, either in terms of their accuracy, or in terms of their tendency to view others as similar to themselves.

These studies have usually been built around a particular operation in which a judge (J) "predicts" how another person (O) will respond. Often, for example, both persons describe themselves on a personality inventory, and J is then asked to fill out the inventory as he thinks O did. The extent to which the prediction agrees with O's actual response is taken as a measure of J's accuracy of social perception (or "empathy", "social sensitivity", or "diagnostic competence", etc). Measurements obtained in this manner are difficult to interpret and several investigators have obtained evidence of distressingly low reliability or consistency for such scores (11,17,19,28).

This paper seeks to disentangle some of the many effects which contribute to social perception scores, and to identify separately measurable components. This analysis has several results:

---

* Appreciation is expressed to Mary E. Ehart, who assisted in all stages of this paper from initial conception to final interpretation; and to Urie Bronfenbrenner and associates, for helpfully providing data and for their courtesy in exchanging ideas throughout our rather similar investigations.

This study was conducted under ONR Contract N6-ori-07135, Fred E. Fiedler, Principal Investigator.

1. It shows that investigators run much risk of giving psychological interpretation to mathematical artifacts, when measures are used which combine the components.
2. It sheds light on the extent to which adaptation to individual differences is advisable, when the differences are not judged accurately.
3. It directs attention to some especially interesting aspects of social perception left untouched by the usual approach.

Our analysis of social perception scores may also be instructive regarding research strategy generally. This area of research has developed in an ultra-operationalist manner; of late, workers have seemed content to regard "empathy" as "what empathy tests measure". The principal research activity has been correlating "empathy", so defined, with other variables. We shall show, however, that the operation involves many unsuspected sources of variation, so that scores are impure and results uninterpretable. Studies based on myopic operationism are largely waste effort, when the operation does not correspond to potentially meaningful constructs. Defining a measure operationally is only a first stage, preliminary to analytic studies which can refine the measure and bring it closer to the intended construct. It may be of interest to note that this paper relies almost entirely on algebraic analysis as its research method. Even though analysis is based on a model rather than actual behavior, it generates several psychological hypotheses.

Although our report deals with a specialized area of perceptual research, it shares much of the perspective of Postman's important general review of perception (24). His remarks are peculiarly pertinent to studies of social perception, even though he was referring especially to the "New Look" studies of perception of words and objects:

"At this juncture of debate, we shall do well to pull up short a moment and reconsider the fundamental operations of our perceptual experiments, particularly as they bear on the validity of the theoretical constructs linking perception to motivation and personality...Experiments have shared a common tendency which may be called the projective bias --- a selective emphasis on central motivational determinants at the expense of adequate attention to the verbal and motor response dispositions of the subject and the relation of these dispositions to the dimensions of the stimulus... We must, then reaffirm the critical importance of a full and precise analysis of the responses as well as the stimuli which furnish the basic data of perceptual experiments."

A Mathematical Resolution of Social Perception Scores into Components

Model and notation

When data are gathered by means of a test consisting of items (i = a, b, c, ... k), these items define a k-dimensional space, and the responses of any person define a point in that space (10). One point is defined by O's actual responses, and another by J's prediction for O. x_{oi} denotes the response given by Other (O) when describing himself on item i. When a judge (J) predicts what O will say, we have a prediction, y_{oij}. x is always used to indicate self-descriptions, y to indicate predictions.

Error in prediction is represented by the discrepancy between x_{oi} and y_{oij}. In some studies J's accuracy is measured by summing

$|y_{oij} - x_{oi}|$. We shall instead measure accuracy by the distance from the predicted to the actual location in k-space, as determined by the sum over items of <u>squared</u> differences. This formula is easier to treat mathematically than the sum of absolute differences, and will ordinarily give about the same results. When all items are of a Yes - No form, so that the error on any prediction can only be 1 or 0, the two formulas give identical results. Our measure has the important property of being invariant under orthogonal rotation of axes (10).

We define the Accuracy with which J perceives O by

$$ACC^2_{oj} = \sum_i (y_{oij} - x_{oi})^2 \qquad (1)$$

Either ACC or ACC^2 may be used as a measure. Because it is much easier to analyze ACC^2 than ACC, we treat ACC^2 throughout this paper. This should be remembered in applying results.

Now we define

$\bar{x}_{..} = \frac{1}{N}\frac{1}{k}\sum_o \sum_i x_{oi}$ — Grand mean of all self-description responses; average elevation).

$\bar{x}_{.i} = \frac{1}{N}\sum_o x_{oi}$ — Mean of all persons on any item

$\bar{x}_{o.} = \frac{1}{k}\sum_i x_{oi}$ — Mean of any person on all items; his elevation

We define $\bar{y}_{..j}$, $\bar{y}_{.ij}$ and $\bar{y}_{o.j}$ correspondingly. Then $\bar{x}_{.i} - \bar{x}_{..}$ would be the mean score on an item, figured as a deviation from the overall mean. $x_{oi} - \bar{x}_{.i}$ is a score on the item, figured as a deviation from the item mean. We shall let x'_{oi} stand for $x_{oi} - \bar{x}_{o.} - \bar{x}_{.i} + \bar{x}_{..}$, and define y'_{oiy} similarly. These are scores figured as deviations from both item and person mean, i.e., "centered on tests and persons."

The Accuracy Score and Its Components

Accuracy in predicting one O on one item may be measured simply by the absolute value of the error. We have the following identity:

$$ACC_{oij} = \left| y_{oij} - x_{oi} \right| = \Big| (\bar{y}_{..j} - \bar{x}_{..}) + ((\bar{y}_{o.j} - \bar{y}_{..j}) - (\bar{x}_{o.} - \bar{x}_{..})) + ((\bar{y}_{.ij} - \bar{y}_{..j}) - (\bar{x}_{.ij} - \bar{x}_{..})) + (y'_{oij} - x'_{oi}) \Big| \qquad (2)$$

In the right member of (2), the first term compares the elevation of J's (the Judge's) predictions to the average elevation of the criterion responses. Each person using a scale forms his own frame of reference, and some people tend to use different portions of the scale than others (7). This first term will be small if J uses about the same part of the scale as the average O.

The second term represents J's ability to predict the relative elevation of the <u>particular</u> O. It measures J's knowledge of this O's mode of response on the scale.

The third term is, like the first, independent of the particular O. It represents J's error in estimating the relative difficulty or popularity of the item.

The fourth component represents accuracy of prediction for the individual Other, after we remove the first three differences. It expresses J's knowledge of the shape and scatter of O's profile of deviation scores.

Most research has been concerned with the accuracy of J as a judge of all Others. This could be represented by an average of his accuracy scores (ACC^2_{oj}) with particular Others.

$$
\begin{aligned}
ACC^2_j = \frac{1}{N}\sum_o ACC^2_{oj} &= k\,(\bar{y}_{..j} - \bar{x}_{..})^2 \\
&+ \frac{k}{N}\sum\left((\bar{y}_{o.j} - \bar{y}_{..j}) - (\bar{x}_{o.} - \bar{x}_{..})\right)^2 \\
&+ \sum_i\left((\bar{y}_{.ij} - \bar{y}_{..j}) - (\bar{x}_{.i} - \bar{x}_{..})\right)^2 \\
&+ \frac{1}{N}\sum_o\sum_i\left((y'_{oij} - x'_{oi})\right)^2 \qquad (3)
\end{aligned}
$$

These four terms are attributable respectively to the difference of J's elevation from the average, his errors in predicting individual deviations in elevation, errors in predicting item means, and errors in predicting individual deviations from the item mean (after correction for elevation). We shall refer to these as the <u>Elevation (E) component</u>, <u>the Differential Elevation (DE) component</u>, <u>the Stereotype Accuracy (SA) component</u>, and <u>the Differential Accuracy (DA) component</u>. The last three of these require separate discussion.

Differential Elevation (DE)

The Differential Elevation component measures J's errors in judging the "elevation" of O's responses. In some tests elevation reflects insignificant response sets, and we should ignore this component (cf. 10, p. 463). In other tests this component reflects J's judgment of the overall "desirability" of each O, and if so, it may be very important.

The Differential Elevation component may be broken down by using the formula for sums of squares of correlated differences:

$$DE_j^2 = k\,(\sigma^2_{\bar{y}_{o.j}} + \sigma^2_{\bar{x}_{o.}} - 2\sigma_{\bar{y}_{o.j}}\,\sigma_{\bar{x}_{o.}}\,r_{\bar{x}_{o.}\,\bar{y}_{o.j}}) \qquad (4)$$

The variance $\sigma^2_{\bar{y}_{o.j}}$ measures J's tendency to predict that Others differ in elevation. It represents Assumed Dispersion in Elevation, later seen to be a component of "Assumed Similarity". $\sigma^2_{\bar{x}_{o.}}$ is the true dispersion in elevation. The correlation $r_{\bar{x}_{o.}\,\bar{y}_{o.j}}$ (to be symbolized DEr) represents J's ability to judge <u>which</u> O's rate highest on the elevation scale. If every item measures morale, for instance, the correlation shows how well J can judge which O's say they have the highest morale.

<u>Stereotype Accuracy (SA)</u>

As used here "stereotype accuracy" refers to the person's ability to predict the norm for Others. It might be called "accuracy in predicting the generalized other" (3). This score depends on J's knowledge of the relative frequency or popularity of the possible responses. In contrast to our stereotype <u>inferred</u> from responses on many items, Gage measures an <u>explicit</u> stereotype. He asks J's to predict the model response among Others of a specified type (16, p. 8-11). Similar stereotype predictions are obtained in studies of ability to estimate group opinion (e.g., 6, 18, 29). Evidence comparing these two types of stereotype would be valuable.

We may write:

$$SA_j^2 = k\,(\sigma^2_{\bar{y}_{.ij}} + \sigma^2_{\bar{y}_{.i}} - 2\,\sigma_{\bar{y}_{.ij}}\,\sigma_{\bar{x}_{.i}}\,r_{\bar{y}_{.ij}\,\bar{x}_{.i}}) \qquad (5)$$

Here each variance is computed over items. The variance $\sigma^2_{\bar{y}_{.ij}}$

expresses how much J expects item means to vary. $\sigma^2_{\bar{x}_{.o}}$ is the scatter of the actual means. SA^2 represents ability to judge the shape and scatter of the profile of item means. $r_{\bar{y}_{.ij}\bar{x}_{.i}}$ (Stereotype Correlation, SAr) represents accuracy in judging mean profile shape without regard to errors in judging profile scatter (i.e., spread in difficulty).

Differential Accuracy (DA)

Differential Accuracy measures ability to predict differences between persons on any item. This component is a sum over items. The component for any item breaks down:

$$DA^2_{ij} = \sigma^2_{y'_{oij}} + \sigma^2_{x'_{oi}} - 2\sigma_{y'_{oij}} \sigma_{x'_{oi}} r_{y'_{oij} x_{oi}} \quad (6)$$

DA^2_{ij}, summed over items, yields DA^2_j. Each variance in the formula is taken over Others. $\sigma_{y'_{oij}}{}^2$ is the Assumed Dispersion on the item (see below). It resembles closely Gage's concept of "rigidity" or "adherence to stereotype" in prediction (15, p. 16; 17).

The correlation (DAr) in (6) is a measure of ability to judge which Others have highest scores on the item, when the score is taken as a deviation from the Others' mean. There is one such correlation for each item.

Implications

Seven aspects of J's performance have been separated:

1. Elevation component: difference of predicted average response from actual average
2. Assumed Dispersion in Elevation } Differential Elevation
3. Elevation Correlation DEr } Differential Elevation
4. Predicted variation in item means } Stereotype Accuracy
5. Stereotype Correlation SAr } Stereotype Accuracy

6. Assumed Dispersion on any item (elevation held constant)
7. Differential Correlation DAr

} Differential Accuracy

The fact that the components are mathematically distinct does not imply that they are necessarily uncorrelated.

Change in any of these may alter the Accuracy score. Surely these aspects of social perception do not all reflect the same trait. A person who uses the same region of the response scale as other persons(Elevation is low) need not have superior insight. And while judging which items have the highest mean seems to require acquaintance with the norms of the group, a person might possess such knowledge to a very high degree and yet lack diagnostic skill which would permit him to differentiate accurately between individuals. At best, failure to recognize the presence of distinct components makes interpretation ambiguous. Chowdry and Newcomb (6) requested group members to predict what percentage of their group would agree with each of many attitude statements. Ability to make this prediction was judged by a difference score, and this score correlated significantly with leadership status. This score, however, combines our Elevation and Stereotype Accuracy components. We cannot conclude that their leaders are better able to judge the specific attitudes in the group. Until the components are separately measured we cannot rule out the possibility that leaders simply used the correct range of the scale more often than non-leaders. This, in turn, might reflect willingness (or unwillingness) to use extreme percentages rather than any more subtle perceptiveness. That such effects do occur is shown in a study by Lorge and Diamond, who required judges to

estimate what proportion of O's would pass ability test items. They found that poor judges were greatly helped simply by being told the difficulty of a few items. "apparently the difference between 'poorest', 'mediocre', and 'best' judges is that the 'best' judges have some experiential reference for the per cent of the population that can pass an item. Giving such referents to the 'poorest' and 'mediocre' judges...leads to a significant reorientation of such judgments." (20,p.33) When judges responded only to the items, the best judges had a mean Stereotype Correlation of .73, and the poorest one of .56. After information indicating an appropriate reference level was given, the same groups had mean correlations of .77 and .73. The difficulty encountered in interpreting the Chowdry-Newcomb study does not arise in two recent treatments of the same problem (18,29) where subjects are asked to predict what ranks will be assigned to certain stimuli. The ranking method eliminates elevation and dispersion differences from the responses, and therefore confines scores to the Stereotype Correlation. An Alternative and more informative method might be to analyze data of the Chowdry-Newcomb type in terms of the separate components so as to determine how leaders behave on each. At worst, failure to identify the components of the Accuracy score leads to artifactual correlations. Only a few of the many examples in the literature need be cited. Norman and Ainsworth (30) report a large number of correlations between Accuracy ("Empathy") and Assumed Similarity ("Projection"). Since the accuracy score contains assumed similarity components, there would necessarily be an overlap between those two scores even in a situation where both sets of responses are determined strictly by chance. The correlations have no psychological meaning. In Dymond's study (12) it was reported that persons with high Accuracy are also most easily judged. But a person who uses the scale in a typical manner will have a low Elevation Component; and thus will have lower Elevation erros in judging him simply because of this typicality. This would happen even if the other predicted his responses without ever meeting him! Perhaps social psychologists should take what comfort they can from Bertrand Russell's remark that physicists "have not yet reached the point where they can distinguish between facts about relativity and mathematical operations which may have nothing to do therewith".

Analysis of Assumed Similarity Score into Components

Assumed Similarity (AS) may be determined (see 14, for example) for a single Other by the formula:

$$AS^2_{jo} = \sum_i (y_{oij} - x_{ji})^2 \qquad (7)$$

y_{oij} is the perception of O by J, and x_{ji} is J's statement about himself. (Sometimes "ASo", Assumed Similarity between two Others selected in a certain manner, is computed). Some investigators have measured AS over many Others, to get a general score called "projection" or "identification" (22).

We may break AS into components as we did ACC. If, as before, we measure AS by a distance formula based on sums of squares,

$$\begin{aligned} AS^2_j = \frac{1}{N} \sum_o \sum_i AS^2_{oij} &= k (\bar{y}_{..j} - \bar{x}_{.j})^2 \\ &+ k \cdot \sigma^2_{\bar{y}_{o.j}} \\ &+ \sum_i ((\bar{y}_{.ij} - \bar{y}_{..j}) - (\bar{x}_{ij} - \bar{x}_{.j}))^2 \\ &+ \sum_i \sigma^2_{y'_{oij}} \qquad (8) \end{aligned}$$

Assumed Similarity, therefore, contains four components. Equation (8) is simpler than the corresponding formula for ACC^2 because some terms vanish.

Assumed Elevation (AE)

The first term we may call the Assumed Elevation (AE) component. It measures J's tendency to assume that Others have the same average response as he does. This component is important if items are polarized so that a high score on each represents good adjustment or some other interpretable quality; the score then shows whether J regards the average O as similar to himself in this central dimension.

Assumed Dispersions (ADE, ADI).

The second component is the Assumed Dispersion among Others in elevation. The fourth is the Assumed Dispersion on specific items after differences in elevation are removed. These dispersions have already been encountered in equations (4) and (6) as components of ACC. We shall refer to them as Assumed Dispersion in Elevation (ADE) and Assumed Dispersion within Items (ADI), respectively.

Assumed Self-Typicality (AST).

The third component measures the discrepancy between J's perception of the average O and his self-description. This component tells whether J regards his own profile as typical in shape. Or, we might say, this component shows the similarity of J's self-perception to his implicit stereotype of Others (elevation held constant). We follow Gage in calling this Assumed Self-Typicality (AST) (16, p.17).

Of the four components, only AST divides into separate variance and correlational terms.

$$AST = k\left(\sigma^2_{\bar{y}_{.ij}} + \sigma^2_{x_{ij}} - 2\,\sigma_{x_{ij}}\,\sigma_{\bar{y}_{.ij}}\,r_{x_{ij}\bar{y}_{.ij}} \right) \qquad (9)$$

The variance among the y's represents the tendency of the predictor to predict different means for different items. The correlation represents the similarity between his self-description and the average profile, after removing differences in elevation and scatter from consideration. We may call it the Self-Typicality Correlation (STr).

To summarize: the components of AS are of two types. ADE and ADI involve Assumed Similarity between Others; i.e., a tendency to differentiate. ARS and AST represent Assumed Similarity of self to average Other. These types seem logically distinct, but a subsequent section will indicate the probable desirability of combining AE with ADE, and AST with ADI.

Optimizing Predictive Decisions

Insofar as our mathematical model is an acceptable approximation to real conditions, we can reason mathematically to determine how a person may improve his judgments. We have assumed that the goodness of predictions can be evaluated by the mean square error. Taking the derivative of each component of ACC^2, and setting that derivative equal to zero, we find that ACC becomes smaller, and therefore prediction improves, when

(a) J has a typical response set.

(b) $\sigma_{\bar{y}_{.ij}}$ approaches $r_{\bar{x}_{.i}\bar{y}_{.ij}} \sigma_{\bar{x}_{.i}}$. Here the variance is over items. This means that $\sigma_{\bar{y}_{.ij}}$ should not exceed $\sigma_{\bar{x}_{.i}}$, and should be near zero if the Stereotype Correlation is low. If this correlation is low, the more he differentiates among items, the poorer is his accuracy.

(c) $\sigma_{y'}$ approaches $r_{x'_i y'_i} \sigma_{x'_i}$, the variance being over Others. This means that $\sigma_{y'}$ should not exceed $\sigma_{x'}$, and should be near zero if the Differential Correlation is low. This principle is true for accuracy of prediction on any single item, and for the elevation score.

It has not been possible to determine the conditions which maximize accuracy as measured by other formulas (such as the mean of ACC_{oj})but a result of the same general character would be expected.

Effects of differentiation on practical decisions

These formal principles indicate that there is an optimal degree of differentiation in making judgments. If a Judge can

make accurate judgments as to the relative location of Others on a continuum, then he is wise to make σ_y as large as σ_x-- never larger. But if he is forced to base his judgment on inadequate cues or if the available personality theory and situational knowledge do not permit trustworthy inference, then he should treat people as if they were very nearly alike. The person who attempts to differentiate individuals on inadequate data introduces error <u>even when the inferences have validity greater than chance</u>. This is consistent with Gage's evidence that judges predict a stranger more correctly when they describe the typical person of his group than when they try to describe him as an individual (16, p.10).

The variation of J's predictions indicates how much he differentiates. For example, a teacher estimating IQ's in a class might spread them from 90 to 110, or from 70 to 130. We would expect the judge who perceives greater differences to apply more sharply differentiated treatments to the persons. His σ_y is essentially a weighting, or an expression of his confidence in his own discriminations (cf. 19, p.201). A person who knows that the expected s.d. for IQ's is 16 might try to predict so that his estimates would have this s.d. but unless he is a perfect judge, this is unwise. He will have less error if his predicted s.d. is less than 16--how much less depending on the correlational accuracy of his predictions.

If two diagnosticians can each judge some trait with correlational validity .40, the one who differentiates strongly (i.e., makes extreme statements) will make far more serious absolute errors than one who differentiates moderately. Indeed, the

person who makes extreme differentiations on the basis of a validity of .40 may make worse predictions, judged by absolute magnitude of errors, than a judge who has zero correlational validity but makes no false differentiations.

"Every pupil has his own pattern of readiness, and the teacher must fit methods to that pattern, not treat the pupil in terms of the statistical average" (9, p.73). Statements such as these, commonly made in teacher-training, now appear to require qualification. From our evidence the degree of adaptation desirable depends on the adequacy of the diagnostic information. If the teacher is not well informed regarding the unique patterns of his pupils, he should probably treat them by a standard pattern of instruction which has been carefully fitted to the typical pupil. Modifying his plans drastically on the basis of limited diagnostic information may do harm. A similar argument applies to clinical diagnosis and industrial leadership. Differentiation is harmful if the extent of adaption or differentiation exceeds the amount justified by our accuracy in differentiating. This is a distinct reversal of the view that judgment is always improved by taking into account additional information which has validity greater than zero. Investigators have noted a "central tendency of judgment", which leads to lower dispersion among estimates than among objects. Whereas formerly "the central tendency of judgment" was regarded as a source of inaccuracy (1, p.521) our analysis shows that this tendency may have beneficial consequences.

Teachers may properly modify treatments considerably to fit individual differences provided they are well able to judge those

differences. They might be expected to judge differences in past achievement in arithmetic quite accurately; if so, they could profitably provide quite different treatments (e.g., different assignments) for different individuals. But if it is hard to judge some other quality (e.g., creative potential in art), then it is a great mistake to differentiate treatment. Treatment from individuals should depart substantially from that suited to the average of the group <u>only</u> when dependable information is available to guide the adaptation.

Illustrative Analysis of Cornell Data

To illustrate our system of analysis, we use data kindly provided by Bronfenbrenner and Dempsey. The data were gathered at Cornell University primarily for the purpose of pilot analyses such as ours. Only eight subjects and nineteen items are involved; and we actually employ only eight of the items.

In the Cornell experiment (4), the eight subjects were candidates for employment as interviewers. Each person interviewed each of the seven others. In each interview, each man was to obtain information about his partner. Following the interview, each person filled out a form stating his own reaction and predicting what his partner would say. There are eight items, each to be judged on a four-point scale. One item is:

To what extent did you feel at ease during the interview?

___a. very much ___b. a good bit ___c. only slightly ___d. not at all

The respective responses are scored 1-2-3-4. Completion of the design provides seven self-descriptions and seven predictions by each man (also, seven <u>for</u> each man).

We have taken two simplifying steps which might be illegitimate for purposes other than demonstration. In every instance, we have used the average of O's responses over all seven interviews as his true response, $x_{oi.}$ This discards information on O's variation from interview to interview. Secondly, we treat J's self-description as a "prediction of himself". This "prediction" is taken as perfectly accurate. By this device, we deal at all times with eight Judges and eight Others, and the criterion is made the same for every person.

Accuracy Scores for Eight Persons

Table 1 presents the ACC score for each person, and his score on each component. Table 2 organizes the same data to show the person's relative position in the group. Based as they are on only eight cases and eight items, these data and subsequent numerical results are illustrative, and not a proper basis for generalization. They may be useful to guide future studies.

Magnitude of components

The Differential Accuracy Component has substantially larger variance than the others, and therefore has much greater influence on ACC. Although Elevation has a smaller mean than Differential Elevation and Stereotype Accuracy, the variancesfor these three scores are nearly equal; they play an equal part in determining individual differences in Accuracy. The correlations DEr and DAr are generally low but positive. The Stereotype Accuracy correlations, however, averaged .74.

Relation of differentiation to accuracy

The data illustrate our mathematical principle that any accuracy component is made smaller as the predicted standard

Table I. Accuracy Scores of Eight Judges Divided into Components

Judge	ACC^2_j	Elevation (E)	Differential Elevation (DE)	Stereotype Component (SA)	Differential Component (DA)	DE contains: $\sigma_{\bar{y}_{o.j}}$	DE contains: $r_{\bar{x}_{o.}\bar{y}_{o.j}}$ (DEr)	SA contains: $\sigma_{\bar{y}_{.ij}}$	SA contains: $r_{\bar{x}_{.i}\bar{y}_{.i.j}}$ (SAr)	DA contains: $\sigma^2_{y'}$	DA contains: $r_{x'y'}$ (DAr)
True $\sigma(\sigma_x)$						.22		.44		.12	
1	2.32	.00	.69	.24	1.40	.24	.18	47	.93	.12	.28
2	5.32	.05	1.71	1.34	2.24	.38	-.14	17	.37	.17	.12
3	4.60	.10	.23	.50	3.77	.15	.63	47	.85	.39	.08
4	5.33	.26	1.35	.91	2.81	.35	.00	21	.67	.25	.23
5	4.37	.98	.79	.28	2.33	.27	.18	45	.91	.16	-.04
6	4.06	1.22	.76	.65	1.45	.18	-.20	54	.86	.11	.26
7	3.75	.20	.73	1.27	1.56	.14	-.44	15	.44	.17	.38
8	3.78	.02	.41	1.28	2.06	.29	.62	76	.92	.27	.35
Mean	4.19	.35	.83	.81	2.20	.25	.10	.40	.74	.20	.21
Variance	.83	.20	.20	.18	.56	.01	.13	.04	.04	.01	.02

Table 2

Accuracy Scores Expressed as Ranks

Judge*	ACC_j^2	Elevation	Differential Elevation	Stereotype Accuracy	Differential Accuracy	Differential Elevation Correlation	Stereotype Accuracy Correlation	Differential Accuracy Correlation
1	1	1	3	1	1	3 1/2	1	3
7	2	5	4	6	3	8	7	1
8	3	2	2	7	4	2	2	2
6	4	8	5	3	2	7	4	4
5	5	7	6	2	6	3 1/2	3	8
3	6	4	1	4	8	1	5	7
2	7	3	8	8	5	6	8	6
4	8	6	7	5	7	5	6	5

* Arranged in order from most to least accurate

deviation (σ_y) approaches the product of the related correlation with the actual standard deviation (σ_x). Consider, for example, the results on Stereotype Accuracy and its constituents. Person 8 is an excellent Judge, according to his stereotype correlation of .92. But he expects too much variation in the item means (.76 compared to an actual s.d. of .44). As a consequence #8 has a poor Stereotype Accuracy score despite his excellent ability to discriminate between items. The best Stereotype Accuracy scores are earned by #1 and #5, who have high correlations and who predicted variance close to the actual variance. Another comparison worth noting involves the Differential Elevation and Differential Elevation Correlation of #3 and #8. These persons have the same Differential Elevation Correlation, but #3 underestimates the variation in elevation, and #8 overestimates it. As expected, #3 earns the better Differential Elevation score.

Reliabilities and Intercorrelations

Reliabilities of components have been determined, where practicable, by the Alpha formula (7). Reliability of Accuracy over items was .14. This value indicates the expected consistency of scores if Accuracy for each predictor were estimated using two independent sets of items. This reliability over Others was .37. This is the consistency expected if Differential Elevation scores were estimated from two independent sets of Others. The results on Differential Components throw further light on this result.

Differential Accuracy was strikingly reliable over items: L = .73. That is to say, some predictors were consistently good

Table 3

Rank Correlations Among Accuracy Components

	Accuracy	Elevation	Differential Elevation	Stereotype Accuracy	Differential Accuracy	Differential Elevation correlations	Stereotype Accuracy correlation	Differential Accuracy Component
Accuracy	(.14)**	.36*	.55*	.33*	.81*	-.06*	.57*	.69*
Elevation			.41	-.21	.35	.44	.26	.33
Differential Elevation			(.37)	.21	.10	.59	.52	.31
Stereotype Accuracy					.24	.21	.64*	-.24
Differential Accuracy					(.73)	-.43	.38	.69*
Differential Elevation correlation							.55	-.29
Stereotype Accuracy correlation								.12
								(.18)

* Not independent

** See text for detailed information on reliabilities

over all items, others consistently poor. But when we examine the components of DA, we find that

(a) σ is reliable over items (αC = .79)

(b) DAr, the measure of accuracy in locating others, is not (αC = .18)

In this sample Differential Accuracy shows reliability because some persons have consistently low assumed similarity. This makes them consistently inaccurate predictors because Differential Accuracy Correlation is generally low). No adequate estimate could be obtained for the reliability of Elevation, of Stereotype Accuracy, or of Accuracy over Others. We examined why Accuracy has reliability much lower than Differential Accuracy, one of its components. Apparently this occurs because the sign of the stereotype error has a substantial effect on accuracy on any one item and therefore lowers the correlation from item to item.

Our limited data suggest the accuracy components tend to be unreliable except where reliable differences in assumed similarity affect the component. Stone and Leavitt (28) report very low consistency (-.07 to 30) of accuracy scores in predicting different children on a fairly long test, but a median consistency of .63 between two halves of the test for the same child. They then trace the latter consistency to consistent favorable sets toward a given child, and to assumed similarity. Further work is needed to establish which independent components of Accuracy can be reliably measured.

In Table 2 we note that Number 1 is consistently superior

on various components of accuracy and #4 is consistently inferior. But #7, the best predictor as judged by Differential Accuracy Correlation is the poorest on Differential Elevation Correlation and next-to-poorest on Stereotype Accuracy Component. Table 3 presents the intercorrelations of the eight measures of accuracy. An asterisk indicates pairs of variates which are experimentally linked; these correlations are higher than would be expected from independent measurements. Being based on only eight cases, the correlations cannot be interpreted confidently. The correlations are low but many of them are as high as the accompanying reliabilities.

Only one firm recommendation can presently be made. Future studies of predictive accuracy should measure the components separately, preferably using two independent sets of items and Others. Such measurement will permit accurate determination of reliabilities, of the relation between the components, and of their relation, if any, to external criteria. Only after such research can we decide how many important components are present within the overall Accuracy score presently used in most research on social perception and which unwanted components must be suppressed by appropriate design of tests and scoring keys.

Illustrative Analysis of Assumed Similarity Scores

In Table 4, the Assumed Similarity scores of the eight Judges are divided into components. Table 5 presents the same information in rank form, and Table 6 presents the intercorrelation. The relatively large variance of ADI, Assumed Dispersion on each item, indicates that it has great influence on individual

Table 4

Assumed Similarity Scores

Judges	Assumed Similarity	Assumed Elevation	Assumed Dispersion in Elevation	Assumed Self-Typicality	Self Typicality Correlation	Assumed Dispersion within Items
1	1.89	.14	.46	.34	.94	.97
2	3.78	.35	1.15	.95	.60	1.36
3	4.63	.06	.18	1.30	.90	3.09
4	3.45	.16	.98	.30	.61	2.01
5	2.21	.01	.59	.37	.89	1.25
6	1.54	.12	.26	.28	.94	.87
7	2.41	.00	.15	.88	.72	1.37
8	4.15	.12	.66	1.16	.89	2.20
Mean	3.01	.12	.55	.70	.81	1.64
Variance	1.14	.23	.12	.16	.02	.49

Table 7

Rank Correlations between

Assumed Similarity Components and Accuracy Components

	Accuracy	Elevation	Differential Elevation	Stereotype Accuracy	Differential Accuracy
Differentiation in Elevation	.55*	-.19	.62*	.38	.24
Differentiation in Pattern	.43*	-.41	.39	.64*	.69

* Not independent

differences in overall Assumed Similarity.

The correlations show great overlap of Assumed Elevation with Assumed Dispersion in Elevation, and Assumed Self-Typicality with Assumed Dispersion within Items. The tendency to differentiate among Others is accompanied by a tendency to differentiate the average Other from oneself. This result is partly an artifact, resulting from using each person's self-description as one of his "predictions". Even allowing for this, our correlation suggests treating only two components of AS: Assumed Similarity in Elevation (AE + ADE) and Assumed Similarity in Pattern (AST + ADI). The correlation between these two variables is negligible (ρ = .21). Further evidence is required to determine how to divide Assumed Similarity and which components merit serious investigation.

Correlation of Assumed Similarity with Accuracy

Table 7 gives the correlations of Assumed Similarity components with Accuracy components.

The Judge's "Implicit Personality Theory"

We turn now to an aspect of social perception data which may prove to be particularly significant. When a Judge makes predictions for a large number of Others, these predictions define a corresponding distribution of points in the variate space. This distribution may be regarded as a description of the generalized Other, representing the Judge's view of both central tendency and individual differences. The Judge's generalized perception may be an important indicator of his expectations regarding others. We shall discuss the general

Table 5

Ranks in Assumed Similarity

Judges*	Assumed Similarity	Assumed Elevation	Assumed Dispersion in Elevation	Assumed Self-Typicality	Self Typicality correlation	Assumed Dispersion within Items
6	1	4 1/2	3	1	2	1
1	2	6	4	3	1	2
5	3	2	5	4	4	3
7	4	1	1	5	6	5
4	5	7	7	2	7	6
2	6	8	8	6	8	4
8	7	4 1/2	6	7	5	7
3	8	3	2	8	3	8

* In order from greatest assumed similarity to least

Table 6

Intercorrelations of Assumed Similarity

		Assumed Similarity	Assumed Elevation	Assumed Dispersion in Elevation	Assumed Self-Typicality	Assumed Dispersion within Items	Self-Typicality correlation
	Assumed Similarity		.08*	.18*	.86*	.93*	.48*
Differentiation on Elevation	Assumed Elevation			.81	-.21	-.11	.28
	Assumed Dispersion in Elevation				-.05	.02	.55
Differentiation on Patterns	Assumed Self-Typicality					.97	.24*
	Assumed Dispersion within Items						.41

* Not independent

significance of this perceptual system before tracing its effect on social perception scores.

J's distribution is to be examined in terms of the means and variances on the original variates, and also in terms of the covariances. The mean may be regarded as J's stereotype; if the mean Other in his descriptions is "hostile", for example, this may be highly significant. The variance indicates J's tendency to differentiate along a dimension; it is represented in Assumed Dispersion within Items. The covariance is interpreted as indicating the relations J expects to find among variates. A given Judge may customarily report the same persons as high on both "quietness" and "shyness", for instance; or on both "ambition" and "selfishness". These aspects of the distribution reveal J's view of others and the connotation of personality traits for him.

We suggest that these means, variances, and covariances describe J's implicit theory of personality. The expectations J has of others constitute his view of personality and presumably direct his responses to Others. While the mean of J's predictions might be consciously controlled to give some impression J regards as desirable, it is quite unlikely that J is aware of the correlation among his responses. For this reason, we believe that the distribution represents J's implicit theory of personality. Certainly it represents associations and norms of which J is not necessarily aware.

Osgood has drawn attention to the possibility of studying the semantic equivalence of stimuli by testing whether they are

used similarly (21). Our method is quite similar, determining as it does what personality traits J regards as occurring together. We propose to examine the reference frame within which J locates particular others.

This concept can be illustrated by using a small portion of the Bronfenbrenner-Dempsey data. Their test required J to predict responses of eight persons (including himself) on these questions:

> 1. In general, how openly did you express your feelings and emotions during the interview?
>
> 2. How much interest did you feel in the other man as a person?
>
> 3. How much were you aware of how he was feeling?
>
> 4. How much opportunity did you give him to interview you?
>
> 5. How much important information were you able to get about him?
>
> 6. To what extent did you feel at ease during the interview?
>
> 7. To what extend did you succeed in establishing a good interviewing relationship?
>
> 8. To what extent did you feel like the person being interviewed rather than the person doing the interviewing?

A four-point response scale was used for these items, a low score indicating a positive answer.

For Judge 3, a poor predictor, we determined the mean, variance, and covariance of his predictions. The matrix of covariances was factored by a pivotal method akin to square-root factor analysis (31), intended to yield interpretable factors. Table 8 shows the loadingson three factors with item means and

variances.

The means for Judge 3 show no striking features, especially when considered in relation to the true means presented below. The variances indicate that #3 regards others as fairly uniform in their awareness of him (item 3), and as varying especially in their openness, ease, and feeling of dominance (items 1,6,8). The first two dimensions of #3's perceptual space are plotted as Figure 1. Little confidence can be placed in factors based on eight cases, but we would otherwise interpret Factor I as representing a feeling of being under pressure. It is notable that #3 regards those persons who are most open (item 1) as being least at ease (item 6). Factor II shows a link between items 4 and 5, getting and giving information. Factor III is indistinct. It is notable that items 6 and 7 are correlated; a "good interviewing relation" is perceived by #3 as one where the <u>interviewer</u> is at ease! Such a finding regarding #3's perception, if better substantiated, might have much diagnostic importance.

The literature contains many studies of correlation between ratings which bear on the perceiver's frame of reference. The studies of halo effect in rating suggest the existence of a strong general good-bad factor. These studies have not examined raters separately. Frenkel-Brunswik reports that ethnocentric individuals see others in black-and-white terms, the "good", "strong" traits going together (15; See also 25). She does not present correlational data, but she is essentially stating that halo effects are stronger in such raters. In our language, their covariance matrix is loaded with one factor, while non-authoritarians use

many factors and do not emphasize the general evaluative dimension. Steiner (27) has substantiated this conclusion, and discusses his results in terms of the perceiver's "trait contingencies", that is, in terms of the perceiver's frame of reference. "The individual's assumption that certain attributes belong together is expected to influence his percept of the person with whom he is interacting" (p.349). Steiner's data are restricted to group differences, but his theory is not. Our position differs slightly from Steiner's in that we emphasize the implicit contingencies of which the perceiver may be quite unaware. Steiner's method, in its present form, requires the perceiver to say explicitly what contingencies he expects.

Two other studies show differences in the perceptual reference frame of groups. Wickman's well known study (30) showed that teachers expected different traits to correlate with mental health than did mental hygienists. Moore (21) performed a factor analysis of ratings given non-commissioned officers by their subordinates, and also of ratings given by their superiors. The factor patterns differed. For instance, superiors coupled "leadership" with eagerness and responsibility, but the subordinates viewed "leadership" as closely linked with intelligence and skill.

None of these studies of groups examines the perceptual space by which an individual describes personality, but the evidence supports the belief that important individual differences exist. In view of our interpretation of the perceptual distribution as an implicit personality theory, special interest

would attach to studies or ratings given by clinical psychologists or psychiatrists of different schools, or having different amounts of training. One objective of instruction in the field of personality is to modify oversimple views students may hold. If our procedure does reveal covert and unconscious conceptions, it may be a useful device for evaluation.

Effect on accuracy scores. The Judge's distribution of Others has been interpreted here as a standing system of meanings which delimits the space within which he locates Others. It is obvious that any such delimitation would affect social perception scores.

Discrepancies between perceived mean and actual mean lower Stereotype Accuracy. We have shown earlier that Accuracy declines if Assumed Dispersion departs from an optimal value. The correlational effects are a bit less easy to perceive.

Correlations describe the shape of the distribution of Others. If traits 1 and 2 are uncorrelated, then Others will have a roughly circular bivariate distribution. If a Judge regards 1 and 2 as correlated, attributing both to the same persons, his perceived distribution will be elliptical. His perceived dispersion along the dimension 1 + 2 will be greater than in the true responses, and his accuracy will suffer. We can view the example in another way. Suppose the judge predicts variate 1 perfectly but believes that variates 1 and 2 correlate 1.00 when they have a true correlation of zero --- then he must have substantial error in predicting variate 2. He can predict 2 accurately only if he perceives the covariance of 1 with 2 accurately.

Data reported by Crow (11, p.86) show this phenomenon clearly. As part of a larger study, he asked Judges to predict what would be the first word missed by a patient on a vocabulary test and what was the highest level attained (called tasks D1 and D2.) The correlation of Judges' accuracy on D1 with accuracy on L2 was positive and significant for five of ten patients, but negative and significant on two patients. Judges tended to expect a correlation between the two scores, and when there was a true correlation they did well; where it was negative, the Judges could not be accurate on both predictions. There was a rank correlation of .97 (over patients) between consistency of accuracy scores, and consistency of the patient's performance.

The Cornell data were examined to determine the covariance between items in self-descriptions. The resulting "criterion" matrix was factored, with the results shown in Table 9 and plotted in Figure 2 (first two factors). This pattern is different from that of #3 (Table 8) in several respects. Notably #3 overdifferentiates on all items. The first factor for #3 lumps openness and lack of receptiveness; these variables are divided among two factors in the criterion. In the criterion, being at ease (item 6) is positively related to openness. It is especially interesting that "feeling like the person being interviewed" is, for the group as a whole, positively correlated with being at ease; but for #3 these items are negatively correlated. With a view of people so discrepant from the facts, it is not surprising that #3 has a low ACC score.

Table 8

Substantial Factors in Judge 3's Covariance Matrix,

Determined by Pivotal Method

Item	I Pressure	II Exchange of Information	III Rapport	h^2	Variance	Mean Prediction
1	2.55	-.68	-.62	7.35	7.58	2.39
2	1.32	-.94	-.16	2.65	3.77	1.80
3	-.76	.16	.82	1.28	1.81	1.70
4	.46	1.35	-.07	2.04	2.34	1.63
5	.68	1.17	.16	1.86	2.16	1.59
6	-1.92	-.76	1.28	5.90	6.42	1.79
7	-.88	-.30	.63	1.26	2.14	1.43
8	2.50	-.08	.40	6.43	6.43	2.93
Sum of Squares	20.21	5.24	3.31	28.77	32.67	
Percent of variance	62%	16%	10%	88%		
Cumulative percent	62%	70%	88%			

Table 9

Factors in the Criterion Covariance Matrix

Determined by Pivotal Method

Item	I Openness	II Receptiveness	III Passivity	h^2	Variance	Mean
1	2.06	-.40	-.30	4.49	4.57	2.05
2	.14	.94	.01	.88	1.03	1.93
3	.09	.96	-.11	.94	1.16	2.04
4	.86	.40	.30	1.00	1.07	1.80
5	.22	.55	-.44	.54	.67	2.05
6	.47	.60	1.16	1.93	1.42	1.57
7	-.10	.86	.09	.76	.97	1.64
8	-.05	-.27	.85	.80	2.45	3.09
Sum of squares	5.29	3.60	2.46	11.35	13.34	16.17 (sum)
Percent of variance	40%	27%	18%	85%		
Cumulative percent	40%	67%	85%			

Recommendations

Studies of perception may be concerned either with constant processes or with variable processes. When social perception is regarded (as in Allport 1, pp. 499-548) as a process of interpreting the expressive cues Other presents, or of empathizing with him, the search is clearly for a variable process. Yet we have seen that the measures currently used are affected by both constant and variable processes, and therefore cannot serve well to investigate either. As Crow states (11, p.57):

> "The difficulty stems from failure to recognize that two meanings of predictive accuracy are involved. The use of the correlation scoring method (either $r_{x_{oi}y_{oij}}$ or $r_{x'_{oi}y'_{oij}}$) defines predictive accuracy as the ability to approximate the actual situation. By the difference score method a subject is penalized for a systematic error in estimation of the magnitude of the actual situation. By the correlation method the subject is not so penalized. Conversely, a subject is penalized by the correlation method if, although he has approximated the actual situation, his predictions do not vary concomitantly with the actual scores. Each of these scoring methods has its advantages and disadvantages. The choice of which technique to use will depend upon the empirical relationship between the procedures."

An argument can be presented for concentrating attention on constant processes, taking up variable processes only after the constant processes are dependably measured. Constant processes in the perceiver have potentially great importance because they affect all his acts of perception. Individual differences in constant processes need to be measured dependably so that their influence can be discounted in studies of variable processes. Moreover, identifying constant errors should permit training to eliminate such biases; this may be the most effective way to improve the social perception of leaders, teachers, and diagnosticians.

Not all constant processes are of theoretical importance. It may be that response sets in filling out a questionnaire, for example, arise as much from the inflection of the experimenter in reading the directions as from any personality characteristic of the subject. It is particularly difficult to decide whether a tendency to say "yes" on questionnaire items, or to give favorable

answers to such questions, is a transient mental set or a reflection of one's personality (7,10). Some decision must be reached, however, in order to design measures which include or exclude these sets as the investigator's theory dictates.

We may venture to suggest which components of social perception measures deserve research attention, recognizing that the ultimate importance of the components depends on whether they relate to important criteria.

(1) To some extent, the Elevation Component reflects whether J interprets the words defining the scale in the same manner as others do. It appears relatively unfruitful, therefore, as a source of information on his perception of others. It should be separately measured or eliminated from consideration.

This is consistent with Postman's view:

> "In experiments concerned with the determinants of perceptual selectivity, the contribution of verbal and motor response habits must be specifically evaluated and wherever possible held constant. The effects of the independent variables can then be evaluated against an empirical baseline defined by the response habits of the subjects."

(2) The Assumed Similarity measures reflect a general orientation toward others. Perhaps the tendency to differentiate which these indices measure is a reaction shown only in the testing situation. But the fact that significant behavioral correlates have been found for Assumed Similarity by Cass, Fiedler, and others (2,5,13,14,25,26) suggests that this is a generalized mental set influencing both test and non-test behavior. Investigators would

do well, however, to consider Postman's conclusion that response dispositions can be established unambiguously only if they are measured by more than one type of response. (24, p.27)

Components related to Assumed Similarity include Assumed Dispersion in Elevation, Assumed Dispersion over items, Assumed Elevation, and Assumed Self-Typicality. Further research is required to determine whether these should be measured separately or combined.

(3) Stereotype Accuracy expresses how closely J's implicit picture of the generalized Other agrees with reality. Differences of this sort are probably important. Attention should be given to the nature of J's errors, as well as to the overall magnitude of the component.

(4) The Judge's Perceptual Space, studied as a whole, includes not only information on his Stereotype and his Assumed Dispersion, but also on the way in which he organizes the field of personality. The only evidence now available on this type of constant cognitive process is sketchy, but it suggests strongly that this is a most important area for research.

(5) The Elevation Correlation and the Differential Accuracy Correlation are measures of J's senstivity to individual differences. It is these measures which will reflect his ability to interpret expressive behavior, or his ability in differential diagnosis. Present evidence on the reliability of measures of this character is not encouraging, and it may be that study of constant processes in social perception will prove more profitable. But those who wish to study "empathy" or "social sensitivity" as it has usually been conceptualized should reduce

their measures to these correlational components. Stereotype components should be eliminated (10, p.458).

Social perception research has been dominated by simple, operationally defined measures. Our analysis has shown that any such measure may combine and thereby conceal important variables, or may depend heavily on unwanted components. Only by careful subdivision of global measures can an investigator hope to know what he is dealing with. Our analysis makes especially clear that the investigator of social perception must develop more explicit theory regarding the constructs he intends to study, so that he can reduce his measures to the genuinely relevant components.

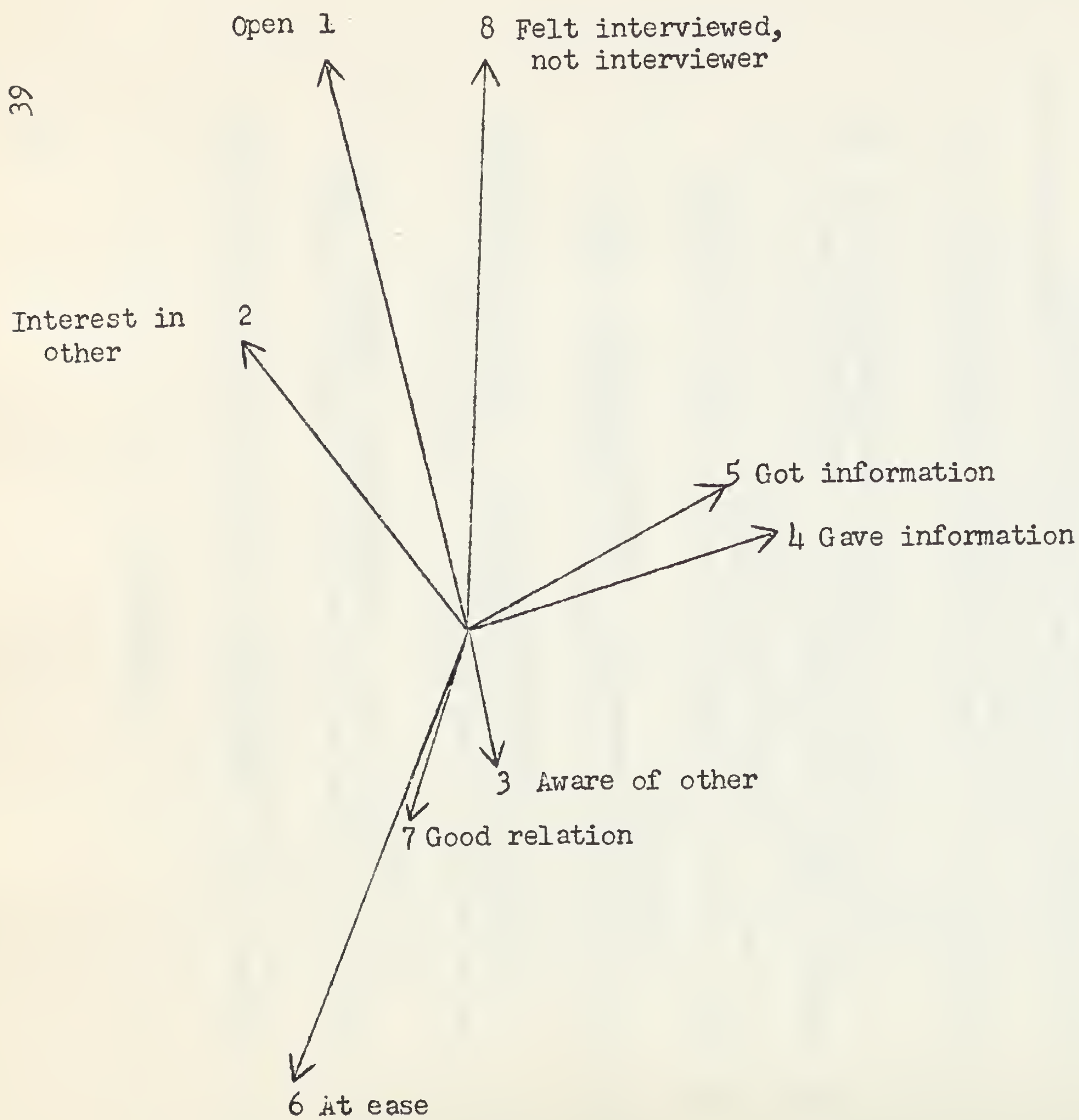

Figure 1. Perceptual space accounting for 78% of #3's perceived variance among Others

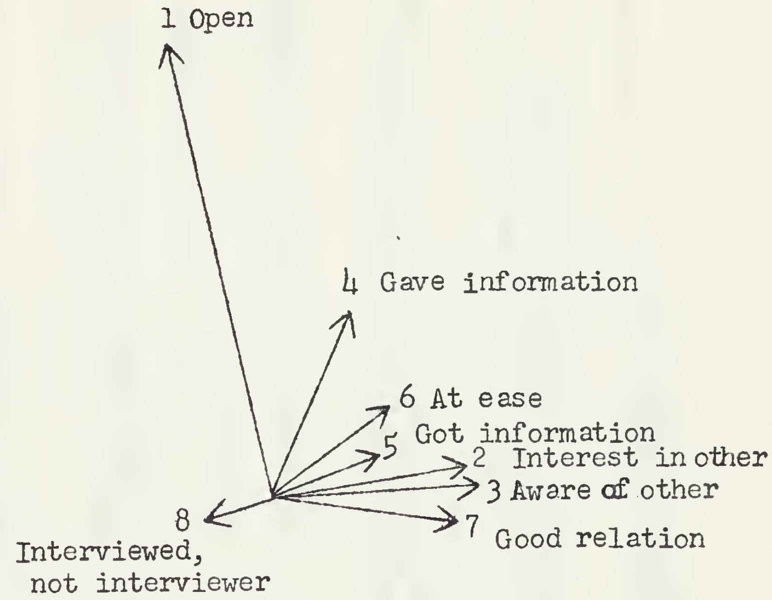

Figure 2. Perceptual space accounting for 67% of the variance in self-descriptions

References

1. Allport, G. W. Personality: a psychological interpretation. New York, Holt, 1937.

2. Bieri, J. Changes in interpersonal perceptions following social interaction. J. abnorm. soc. Psychol., 1953, 48, 61-66.

3. Bronfenbrenner, U. and Harding, J. Components and correlates of social sensitivity. Amer. Psychologist, 1954, 9, in press (Abstract).

4. Bronfenbrenner, U. and others. The analysis of social sensitivity (empathy). Amer. Psychologist, 1952, 7 (Abstract).

5. Cass, Loretta K. Parent-child relationships and delinquency. J. abnorm. soc. Psychol., 1952, 47, 101-104.

6. Chowdry, Kalma, and Newcomb, T. M. The relative abilities of leaders and non-leaders to estimate opinions of their own groups. J. abnorm. soc. Psychol., 1952, 47, 51-57.

7. Cronbach, L. J. Further evidence on response sets and test design. Educ. psychol. Measmt. 1950, 10, 3-31.

8. Cronbach, L. J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.

9. Cronbach, L. J. Educational psychology. New York: Harcourt, Brace, 1954.

10. Cronbach, L. J., and Gleser, Goldine C. Assessing similarity between profiles. Psychol. Bull., 1953, 50, 456-473.

11. Crow, W. J. A methodological study of social perceptiveness. Ph. D. Thesis, U. Colorado, 1954.

12. Dymond, Rosalind F. Personality and empathy. J. consulting psychol. 1950, 14, 343-350.

13. Fiedler, F. E. A method of objective quantification of certain counter-transference attitudes. J. clin. Psychol., 1951, 7, 101-107.

14. Fiedler, F. E. Assumed similarity measures as predictors of team effectiveness. J. abnorm. soc. Psychol., 1954, 49, 381-388.

15. Frenkel-Brunswik, Else. A study of prejudice in children. Human Relations, 1948, 1, 295-306.

16. Gage, N. L. Judging interests from expressive behavior. Psychol. Monogr., 1952, 66, No. 18 (Whole No. 350).

17. Gage, N. L. Accuracy of social perception and effectiveness in interpersonal relationships. J. Person., 1953, 22, 128-141.

18. Greer, L. F., Galanter, E. H., and Nordlie, P. G. Interpersonal knowledge and individual and group effectiveness. J. abnorm. soc. Psychol., 1954, 49, 411-414.

19. Kelly, E. L. and Fiske, D. W. The prediction of performance in clinical psychology. Ann Arbor: Univer. of Michigan Press, 1951.

20. Lorge, I. and Diamond, Lorraine K. The value of information to good and poor judges of item difficulty. Ed. Psychol. Measmt., 1954, 14, 29-33.

21. Moore, J. V. Factor analytic comparisons of superior and subordinate ratings of the same NCO supervisors. HRRC Technical Report 53-24, July, 1953.

22. Norman, R. D. and Ainsworth, Patricia. The relationships among projection, empathy, reality, and adjustment, operationally defined. J. consult. Psychol, 1954, 18, 53-58.

23. Osgood, C. E. The nature and measurement of meaning. Psychol. Bull., 1952, 49, 197-237.

24. Postman, L. Perception, motivation and behavior. J. Person., 1953, 22, 17-31.

25. Scodel, A. and Mussen, P. Social perceptions of authoritarians and nonauthoritarians. J. abnorm. soc. Psychol., 1953, 48, 181-184.

26. Sopchak, A. L. Parential "identification" and "tendency toward disorders" as measured by the Minnesota Multiphasic Personality Inventory. J. abnorm. soc. Psychol., 1952, 47, 159-165.

27. Steiner, I. D. Ethnocentrism and tolerance of trait "inconsistency". J. abnorm. soc. Psychol., 1954, 49, 349-354.

28. Stone, G. C. and Leavitt, G. S. Generality of accuracy in perceiving standard persons. Paper presented to Midwestern Psychological Association, May, 1954.

29. Talland, G. A. The assessment of group opinion by leaders, and their influence on its formation. J. abnorm. soc. Psychol., 1954, 49, 431-434.

30. Wickman, E. K. Children's behavior and teachers' attitudes. New York: Commonwealth Fund, 1928.

31. Wrigley, C. and McQuitty L. L. The square root method of factor analysis: a re-examination and a shortened procedure. Mimeographed, Univer. of Illinois, 1953.