

Individual Differences Correlates of Accuracy in Evaluating Others' Performance Effectiveness

Walter C. Borman

Personnel Decisions Research Institute

This paper presents an approach to studying accuracy in person perception. Problems in assessing interpersonal accuracy are discussed, and then addressed in an empirical study of the perception of human performance. In the study, 16 scripts describing persons performing on two jobs—recruiting interviewer and manager—were prepared in such a way that the performers' effectiveness on various dimensions of performance approached a preset, realistic level. Five- to nine-minute performances of these scripts were videotaped, and "true scores" of effectiveness were developed by obtaining expert ratings of performance on each relevant job dimension. One hundred forty-six students then completed a series of inventories tapping various individual differences and rated the performers' effectiveness on each dimension. Differential accuracy (DA; Cronbach, 1955) scores were computed for each subject, and DA scores were correlated with inventory responses. Seventeen percent of the variance in DA was accounted for by the individual differences measured. A profile of the accurate perceiver of performance was developed, and these results were compared to findings from other studies of interpersonal accuracy.

Accuracy in person perception has proven difficult to study. Several methodological and conceptual problems have slowed research advances in this area, though the study of the person perception *process* continues to draw much atten-

tion (e.g., Tagiuri, 1969). Briefly, three of the most critical difficulties in studying interpersonal accuracy are (1) criterion problems related to establishing justifiable "true scores" reflecting each target ratee's true standing on the traits or dimensions of interest; (2) methodological pitfalls in selecting scoring indices for determining a rater's accuracy; and (3) instability of accuracy "scores" for individuals across different situations. The purpose of this paper is to discuss these difficulties and to present the results of an empirical study that represents an approach to studying interpersonal accuracy within certain important contexts, an approach that avoids some of the problems encountered in this area of investigation. The study also examines individual differences correlates of accuracy in one such context—accuracy in evaluating others' performance effectiveness.

Criterion Problems

Criterion problems in interpersonal perception relate to the question of how to assess a person's accuracy when no solid, justifiable criterion measures are available. Criterion problems are especially severe when dealing with such abstract concepts as personality traits. For example, regarding "dominance," what final authority can decide whether peer ratings, self ratings, or personality test scores should be used

as *the* measure of dominance in the assignment of “true” criterion scores to target ratees? All three of these measurement methods have conceptual strengths and weaknesses. Likewise, other traits or concepts not directly and objectively scorable are subject to this kind of criterion problem.

Problems with Statistical Methods of Measuring Interpersonal Accuracy

This set of problems relates to the appropriateness of various statistical procedures for assessing accuracy in person perception. Briefly, Cronbach (1955) severely criticized the summary D^2 statistic often used in person perception studies to index accuracy. He showed that a measure of D^2 between a rater's ratings and the criterion scores contained four *different* accuracy components: elevation, differential elevation, stereotype accuracy, and differential accuracy. Further, he speculated that these different components might yield relatively uncorrelated scores.

Cline (1964) later demonstrated empirically that these scores did, in fact, intercorrelate only minimally. This result suggests that a generalized “ability” to rate others accurately does not exist; persons may appear to be good at perceiving others when accuracy is defined one way, but seem poor at perceiving the same individuals when accuracy is defined in another way. Thus, even if criterion true scores could be assigned without error, there exist difficulties in selecting an appropriate strategy for computing accuracy scores.

Instability of Individuals' Interpersonal Accuracy in Different Settings

Suppose that it is possible to assign justifiable true scores to target ratees on some traits or constructs of interest, and that there was a single appropriate way to measure perceptual accuracy. The interpersonal accuracy individuals display may still depend largely on the situation, with persons perceiving others relatively accurately in one setting and not so accurately in

other settings. Gage and Cronbach (1955), Taft (1955), Cline (1964), and Dunnette (1968) have outlined large numbers of contexts in which persons might be expected to exercise person perception skills. For example, the form of interaction between perceiver and target ratee may vary widely. In addition, the perceiver might be required to rate the personal characteristics or to predict the future behavior of a ratee, or to make some other judgment or prediction/post-diction about the ratee. Further, the information about a ratee available to the perceiver might be considerable or meager, relevant or irrelevant, and so forth. In short, the conditions under which individuals perceive and make judgments about others vary greatly. Each of these conditions may require different skills and abilities for the accurate perception of others.

Addressing These Problems: General Description of the Study

First, with respect to criterion problems, one approach to “attacking” these problems is to define very carefully the constructs of interest, and then to secure multiple measures of each construct for each target person, measures that provide conceptually appropriate, independently derived estimates of criterion scores. If good across-measure agreement is obtained for these kinds of scores, there is some justification for using them as “true score” estimates.

Second, the problems related to the statistical procedures appropriate for assessing accuracy in person perceptions may not be so complex and hopeless if the psychological significance of each of the various accuracy indices is examined closely. Regarding Cronbach's four accuracy components, in most situations *differential accuracy* appears to be the only conceptually appropriate index. Typically, we are concerned with how accurately the rater can rank order target persons on some dimension. For example, we might want to know how correctly a person evaluates a number of target ratees on intelligence. Criterion problems aside, the index applicable

in this setting is differential accuracy—the correlation between ratings of each target person on intelligence and their “true scores” on this construct.

Others have also argued for using differential accuracy as *the* measure of perceptual accuracy. Sechrest and Jackson (1961) state, “. . . the problem of eliminating response biases is not insurmountable. The solution seems to lie in requiring differential predictions across objects” (p. 168). And Hastorf, Schneider, and Polefka (1970) write, “This [differential accuracy] score is probably closest to what a sophisticated reader of the research literature in this [person perception] area would regard as a ‘true’ accuracy score because the various ‘response bias’ components [elevation, differential elevation, and stereotype accuracy] have been eliminated” (p. 32).

Finally, with respect to the third problem, it is likely that certain people are good at perceiving others in some contexts, while others are more accurate in making person perception judgments in other settings. Gage and Cronbach (1955) adopted this position and argued that it should come as no surprise that little correlation exists among accuracy scores obtained in very different settings. Indeed, studies have found that individuals’ accuracy scores across situations are not highly correlated (Cline, 1964; Crow & Hammond, 1957).¹ This suggests that individual differences associated with person

perception accuracy may be different for different situations. However, within a relatively narrow range of contexts, across-situation interpersonal accuracy may be a more consistent phenomenon.

Thus, one approach which may avoid the generality problem is to study accuracy within more narrowly defined ranges of situations. If certain situations that are of particular interest can be identified, accuracy can be studied in depth within those contexts. Then, within these relatively homogeneous contexts, individual differences associated with this kind of interpersonal accuracy “ability” can also be studied.

The study described here addresses all three of the issues discussed above. First, the third issue was addressed by choosing to study one particular aspect of interpersonal accuracy, the perception of human performance (i.e., rather than the perception of others’ personality or the prediction of others’ behavior). Using this approach, it was hoped to study one important facet of interpersonal accuracy, a facet narrow enough that person perception abilities would be stable within the limits of the context defined.

Second, the criterion problem was addressed by carefully defining the performance constructs to be rated by perceiver subjects and by obtaining multiple expert judgments of criterion performance scores for ratees on these constructs. Convergent and discriminant validities (Kavanagh, MacKinney, & Wolins, 1971) of these expert ratings were then assessed to determine the “goodness” of the criterion performance scores. And finally, Cronbach’s differential accuracy (DA) index of interpersonal accuracy was employed in the study.

Method

Summary of Procedures

Behaviorally anchored rating scales (Smith & Kendall, 1963) were developed for the jobs of recruiting interviewer and manager in a problem-solving session. Then scripts describing persons performing on the two jobs were prepared in

¹In reviewing the interpersonal accuracy literature, Taft (1955), Allport (1937), and Bruner and Tagiuri (1954) concluded that certain personal characteristics appear to describe the “good judge” over a variety of settings. However, Cronbach’s (1955) and Crow and Hammond’s (1957) more recent work suggest that at least some of the generality in the accuracy “ability” was caused by artifactual across-situation relationships. Crow and Hammond (1957), for example, found that rating response tendencies were consistent across situations but that when these response biases were removed and a pure differential accuracy measure employed, reliability dropped to near zero. They concluded that response characteristics (such as leniency and restriction of range) imbedded in the earlier accuracy measures may well have led to the consistency in accuracy scores across situations.

such a way that the performers' effectiveness on various dimensions of performance approached a preset level. Scripts were written depicting five- to nine-minute performances, with the recruiters and managers in these scripts reflecting as closely as possible the preset performance levels. Videotapes were prepared using the scripts, and performance ratings from 14 experts were obtained. Because of the high convergent and discriminant validities obtained from the expert ratings, and due to high correlations between their judgments and *intended* true scores, their mean ratings were chosen to serve as the criteria against which subjects' ratings were compared.

One hundred forty-six college students then completed a series of inventories and used the same behavior scales to rate the effectiveness of persons depicted in the videotapes. Rating accuracy was determined by using Cronbach's DA measure, and students' scores on the inventory scales were correlated with DA. Procedures for developing criterion "true" scores for ratees and accuracy scores for subjects are described in more detail in Borman, Hough & Dunnette (1976).

Rating Scales

Six 7-point behavior-oriented rating scales (Smith & Kendall, 1963) were developed for the

recruiter job. They were (1) creating a favorable image for the company; (2) organizing the interview; (3) providing relevant information about the company; (4) asking relevant questions; (5) answering recruits' questions; and (6) establishing rapport with recruits. Seven 7-point scales were developed for the job of manager in a problem-solving session. (See Table 1 for the dimension labels.)

Generating "Intended True Scores" for Performers

To make performances on each job as realistic as possible, "intended true scores" with a preset covariance structure were generated for the performances related to each job. The two realistic covariance matrices were formed first by asking five experts to estimate the "true" intercorrelations among dimensions; when high reliability was obtained for that task (intraclass r 's = .82 and .81 for the two jobs), and after means of 4.0 and standard deviations of 1.5 were assigned to each dimension, performance profiles reflecting the "correct" covariance structures were generated for ratees performing on the two jobs following a method outlined by Naylor and Wherry (1965). As an example, Table 1 shows intended performance profiles for two ratees on the manager job. The resulting matrices, then, provided realistic multidimensional per-

Table 1
Performance Profiles Reflecting Intended
True Scores for Two Managers

Performance Dimension	Profile 1	Profile 2
Structuring and Controlling the Interview	6.0	2.0
Establishing and Maintaining Rapport	4.5	6.0
Reacting to Stress	5.0	3.5
Obtaining Information	6.0	2.5
Resolving Conflict	6.0	5.0
Developing the Subordinate	3.5	3.5
Motivating the Subordinate	5.0	2.5

formance profiles for 16 individuals—8 on the manager job and 8 on the recruiter job.

Developing and Videotaping the Performance

Sixteen scripts were written, each depicting either a manager talking with his/her problem subordinate or a recruiter interviewing a prospective employee. The scripts reflected as closely as possible the performance levels defined by the intended true scores. Actors were selected to play each role. The same actor played the recruitee in all eight recruiter performances, and another actor played the disgruntled employee in all eight manager performances. Sixteen different actors played the various recruiter and manager roles. Each actor conformed closely to the script during the videotaping session.

Obtaining Final True Scores

Fourteen expert raters were selected to evaluate the effectiveness of each performer. Seven of the raters were graduate students in counseling or industrial psychology. The other seven raters were practicing industrial psychologists working either for a psychological consulting firm or in the personnel research department of a large manufacturing company. All raters were very familiar with the performance demands of the two jobs. Some of these expert judges wrote scripts or acted in the videotaped performances; but none knew the means, standard deviations, or intercorrelations of the intended true scores, nor did they rate their own performances.

Scripts for the performances were revised as necessary to reflect the verbal behavior actually depicted in the performances, and raters were asked to study these scripts and the rating scales before coming to the rating sessions, since it was desired that these expert raters be extremely well prepared to make the performance ratings. They were urged to become thoroughly familiar with each ratee's performance. The rating ses-

sions themselves provided ample opportunity to review notes and to think carefully about ratings. It was hoped that maximum opportunity to review relevant performance-related behavior would lead to highly informed and valid expert ratings.

Experts' ratings were analyzed using an indirect validation approach. Convergent and discriminant validities were assessed using Kavanagh, MacKinney, and Wolins' (1971) ANOVA analysis. Convergent validities (ratee effect in Kavanagh et al.'s design) for the two jobs were highly significant; and the intraclass indices suggested for across-study comparisons were as high as, if not higher than, those obtained in other studies which provide such indices (Borman, 1978). More impressive were the discriminant validities (ratee \times dimension interaction) associated with these expert ratings. Intraclass indices for the two jobs were .57 and .58, with the lower bounds of the .05 confidence intervals equal to .42 and .44. The highest index obtained in the other studies reviewed was .40, with a median index of .15 for these studies. To a considerable extent, expert raters in this study thus reliably differentiated ratee performance on the different dimensions. Further, correlations between mean expert ratings and intended true scores for individual dimensions were all above .65 except for one ($r = .42$), with the median $r = .91$. Therefore, mean expert ratings were adopted as the "true scores" for subsequently assessing raters' differential accuracy.

Developing the "Predictor" Variable Set

The Minnesota Person Perception Battery (MPPB) was developed to tap personality, interest, ability, and life history variables which a literature review suggested might be related to interpersonal perception ability and, more specifically, to the ability to rate human performance. Table 2 contains a list of the measures included and descriptions of scales for which no published information is available. Some of the measures were derived from pub-

Table 2
Description of the Minnesota
Person Perception Battery

-
1. Biographical and Preferences Inventory
Twenty-four items targeted toward five categories;
a) effort expended on tasks, attention span, problem solving style; b) past and present sociability, leadership experience; c) experience in rating others; d) detail orientation; e) demographic information (sex, age, etc.)
 2. Self-Monitoring Scale (Snyder, 1974)
 3. Kelly Reperatory Grid (Kelly, 1955)
Measure of cognitive complexity and "least preferred person" were derived, the latter measure similar in concept to Fiedler's LPC scale (Fiedler, 1964)
 4. Self-ratings on eight evaluative dimensions
Eight bi-polar rating scales (e.g., cheerful - ill-humored; outgoing - shy)
 5. California Psychological Inventory (Tolerance, Sense of Well Being, Social Presence - Gough, 1957)
 6. Empathy Scale (Hogan, 1969)
 7. Estimate of Self-Competence Scale (Motowidlo, in press)
 8. Gough's Opinion Prediction Scale
Forty items requiring estimates of percent of people - in-general responding to opinion-type statements
 9. Wesman Personnel Classification Test
Forty items measuring verbal reasoning and analytical skills
 10. Differential Personality Questionnaire (Stress, Authoritarianism, Hark Work, Impulsiveness, Absorption - Tellegen, 1975)
 11. Janz Person Perception Inventory
Fourteen items developed to emasure two dimensions: seeing other persons as similar to each other; affect related to evaluating others. These dimensions emerged in our factor analytic studies of a larger number of items intended to tap constructs hypothesized to relate to interpersonal accuracy
 12. Self ranking on the six Holland Interest Themes (Holland, Whitney, Cole & Richards, 1969)
 13. Adjective Check List (Gough & Heilbrun, 1965)
-

lished tests or inventories, but several were developed especially for this research.

Data Collection Procedures

Two hundred fifty-eight persons volunteered for this person perception experiment. All but three were students at the University of Minnesota. Subjects attended the first of three sessions in groups of 26 to 60. The same experimenter administered the MPPB to all groups. Then MPPB scores were used to select persons for the performance rating sessions. This was accomplished by employing a χ^2 procedure² to rank order subjects according to their MPPB scores' deviations from the total groups' mean scores on the several MPPB variables. The 150 persons with the highest χ^2 scores were subsequently invited to the performance rating sessions.

One hundred forty-six subjects did attend the rating sessions in groups of 7 to 18; the usual size was 13 to 15 per session. Each subject attended two sessions: one in which he/she viewed recruiter tapes and the other in which he/she viewed manager tapes. The order of viewing jobs was counterbalanced; and within each session, the eight taped performances were presented in random order.

At the beginning of each session, the administrator explained the duties of the job which was about to be viewed and talked about each of the rating scales. She stated that the performance requirements for the job of manager or recruiter were defined by the content of the rating scales and that effectiveness should be judged by referring carefully to these performance dimensions. She explained how to use the behavior-based scales and described the procedures to be followed during the rating session.

After the explanations were completed, subjects viewed the first tape and rated the performer—a recruiter or manager—on all scales relevant to the particular job being observed.

Next, the subjects viewed the second tape and rated the performer's effectiveness on the rating scales. This procedure was followed for all eight job performances. Approximately one week later, subjects returned to view the taped performances of the other job. The administrator followed the same procedure during the second rating session.

Developing Predictor Composites from Responses to the MPPB

The goal here was to represent as comprehensively and as parsimoniously as possible a wide range of MPPB variables. Thus, a factor analysis was employed to form composites, each serving as a summary construct of several variables measured by the MPPB. Although the relationships between individual items/scales on the MPPB and various rating data were also subsequently examined, the composites allowed for a summary of these relationships.

Fifty-seven variables were intercorrelated and factor analyzed (using the $N=258$ sample). Two to 20 factors were extracted, with each solution rotated to the varimax criterion. The 18-factor solution was selected as most psychologically meaningful. Fourteen of these 18 factors were interpretable and were used in subsequent analyses.

Since the intent here was to summarize MPPB variables and not necessarily to eliminate measures, the 18-factor solution was examined for variables that were reliably measured but not well represented within any of the 14 factors defined above. Two measures were identified in this manner: (1) Gough's (1957) 40-item scale tapping accuracy in perceiving the opinions of people in general and (2) the LPP measure derived from the Kelly (1955) Repertory Grid ratings. Both measures possessed good reliability and correlated minimally with the factors; therefore, these two variables were added to the summary predictor set. The resulting 16 variates accounted for approximately 60% of the variance in the total 57×57 correlation matrix

²Write the author for details of this multivariate procedure.

and reflected a parsimonious and nearly orthogonal summary of the 57 variables.

Data Analysis Procedures

Data analyses investigating relationships between individual differences and rating accuracy proceeded in three steps. First studied was the consistency of rating accuracy across the two rating tasks, i.e., across the two jobs. High consistency indicates that subjects accurate in rating performance on one job tended to be accurate also in their ratings on the other job; low consistency means that subjects' accuracy in ratings on one job had little to do with their accuracy in ratings on the other job. Intraclass correlation was employed to estimate the reliability (consistency) of the pooled, across-job accuracy scores.

The second data analysis step was to regress MPPB composite measures upon DA for each job separately and for the two jobs combined. Accordingly, factor scores were computed for each subject on each of the 14 factors, and standard scores on the other 2 variates were calculated for each subject. Then a multiple regression was employed to evaluate relationships between the 16 variates and accuracy. This multiple regression analysis provided a multivariate depiction of the predictability of performance rating accuracy from the MPPB variates and also a view of the stability of relationships between individual differences and interpersonal accuracy.

In the third analysis, the relationships between individual MPPB measures and accuracy were examined to characterize in more detail the "accurate perceiver of performance effectiveness." Again, analyses were performed for each job separately and for the jobs combined to provide estimates of the stability of these results.

Results

The across-job consistency of DA scores for individuals in the sample was moderate (intra-

class $r = .46$), indicating that in the two situations studied (i.e., the two different jobs) interpersonal accuracy was a reasonably stable (within-rater) phenomenon. This stability was considered to be sufficiently high for pursuing an investigation of individual differences-interpersonal accuracy relationships within this context (Borman, 1977).

As mentioned, the 16 composite variates were regressed upon DA scores for each job separately and for the two jobs combined. Zero-order correlations and multiple R s appear in Table 3. These coefficients—.44, .47, and .52—probably overestimate the magnitude of the population multiple R s. Therefore, Wherry estimates (Wherry, 1931) were employed to estimate these population R s. The resulting R s were .28, .33, and .41, respectively, for the recruiter job, the manager job, and the two jobs combined.

Univariate r 's between individual composites and accuracy show that Variables 3, 11, and 12 in Table 3 were consistently related to accuracy in the positive direction.³ Accurate raters in this study tended to be free from self-doubt, tended not to worry or to become stressed, were intelligent, had high grades and investigative interests, and tended to be detail oriented in their approach to tasks.

Correlations between individual MPPB variates and DA were also examined to help characterize the accurate perceiver of performance. Table 4 presents the 10 variables related significantly to DA when accuracy data from both jobs were combined to form the criterion.

³Actually, these univariate r 's may provide underestimates of the relationships between these constructs and accuracy because the factor reliabilities were certainly less than 1.0 and perhaps quite low, especially for the last few factors extracted. Unfortunately, it was not possible to generate reasonable internal consistency estimates of these reliabilities; and, therefore, the extent of the restriction in validities due to unreliability of the factors is unknown.

Table 3
Multiple Regression Results: Regressing MPPB Variates on Accuracy

	Univariate Correlations ^a		
	Recruiter Job	Manager Job	Both Jobs Together
1. Friendliness; consideration; interpersonal orientation	.02	-.07	-.04
2. Organization; hard work	.08	.02	.06
3. Freedom from self-doubt and disillusionment; acceptance of others; tendency not to worry or become stressed	.20	.21	.26
4. Dominance, decisiveness, and perception of self-competence; achievement motivation	-.05	-.05	-.06
5. Conventional versus artistic interests	-.15	-.06	-.12
6. "Old age"; experience in evaluating others' performance as well as interest and involvement in the process	-.04	-.02	-.03
7. Preference for dealing with things rather than people; realistic versus enterprising interests	-.03	-.02	-.03
8. Youngest in family	-.06	-.05	-.07
9. Social interests; interest in observing others	.00	-.24	-.17
10. Sensitivity and social responsiveness; being interpersonally attuned; impulsiveness	.13	.02	.08
11. Intelligence, high grades, and investigative interests	.20	.21	.26
12. Detail orientation	.16	.12	.18
13. Task involvement; endorsement of strict discipline	.05	.14	.13
14. Number of friends and leadership in high school; self-perceived similarity to others	-.03	-.11	-.09
15. Tendency to evaluate favorably those persons not liked	.05	.04	.05
16. Knowledge of the general public's opinions	.09	.03	.07
Multiple correlations	.44	.47	.52
Wherry estimates	.28	.33	.41

^aRegression weights are not provided in this table because they are very similar in magnitude to the univariate correlations due to the near orthogonality of the 16 variables.

Table 4
Highest Correlates of Accuracy

N	Variable	Correlation with Overall Accuracy	Correlation with Accuracy on Recruiter Job	Correlation with Accuracy on Manager Job
144	1. Verbal reasoning	.31**	.21*	.28**
145	2. Self-perceived detail orientation--Bio items	.24**	.25**	.14
145	3. Number of unfavorable adjectives checked--ACL	-.20*	-.10	-.21*
145	4. Self-control--ACL	.20*	.13	.19*
143	5. Tolerance--CPI	.19*	.20*	.11
145	6. Heterosexuality--ACL	-.18*	-.16	-.13
145	7. Exhibition--ACL	-.18*	-.20*	-.10
145	8. Self-perception of how often first impression of a person is right--Bio item	.17*	.10	.16
141	9. Empathy--CPI, MMPI	.17*	.19*	.09
145	10. Aggression--ACL	-.17*	-.09	-.17*

* $p < .05$; ** $p < .01$

Discussion

Within this study, individual differences related to personality, interests, ability, and background explained approximately 17% of the variance in total performance rating accuracy according to the Wherry estimate of population R . This result suggests that individual differences probably play a significant role in determining a person's accuracy in evaluating others' performance.

Further, the study yielded the following characterization of the accurate performance rater, though it should be emphasized that the predictor-accuracy relationships were not very large. Adjective Check List and California Psychological Inventory correlates suggested that accurate raters tended to be dependable, stable, good-natured persons and were seldom described in such terms as rebellious, arrogant, careless, headstrong, irresponsible, disorderly, or impulsive (Gough & Heilbrun, 1965). Positive correlations with "Empathy" and "Tolerance" helped to complete the description of the accurate perceiver. Using results of Hogan's research on the empathy construct and Gough's work on the tolerance construct, the data tentatively suggest that the accurate perceiver of performance is generally even-tempered, outgoing, patient, affiliative, but socially ascendant (Grief & Hogan, 1973). Accurate perceivers also tend to be informal, pleasant, logical, unselfish, and mature. They are verbally fluent and conversationally facile, usually taking the initiative in social relations (Gough, 1968).

Again, the correlations on which this characterization is based were in general low; therefore, this profile should be considered tentative. However, most of the relationships between individual MPPB measures and accuracy were reasonably consistent across jobs, providing some support for the stability of these relationships.

The generality of these results should, however, be tested in other kinds of interpersonal perception contexts. A specific-general hypothesis of the relationship between individual differences and accuracy formulated by Borman,

Hough, and Dunnette (1976) postulates that certain individual differences may be consistently related to interpersonal accuracy across a wide range of situations, while other individual differences may correlate with person perception accuracy in only specific kinds of contexts. Provided that ways can be found to study interpersonal accuracy in other contexts—ways that overcome the difficulties discussed previously in this paper—the generality of these relationships between individual differences and interpersonal accuracy should be investigated both in various performance evaluation contexts and in contexts related to making other kinds of interpersonal judgments.

A curious similarity between the perceiver individual differences found related to interpersonal accuracy in this study and the perceiver characteristics identified as correlates of accuracy in other studies (see especially Taft's 1955 review) lends support to the hypothesis that at least some perceiver characteristics are important for interpersonal accuracy in a variety of settings. Table 5 presents individual differences correlates of accuracy derived from previous studies and also presents the significant ($p < .05$) correlates of DA from *this* study, both composites (as listed in Table 3) and individual MPPB measures being included. Thus, even though severe methodological difficulties are inherent in most of the previous studies cited in the table, especially problems with the D or D^2 index of accuracy (Cronbach, 1955), the patterns of correlations for those studies are very similar to the pattern of relationships found in the present study. Possibly, the measurement problems in previous studies, though severe from a conceptual point of view (e.g., Gage & Cronbach, 1955), do not make much difference in practice. Cline's (1964) finding that D^2 correlated only .25 with DA in one of his studies argues against this interpretation, however. Again, the generality of these individual differences-interpersonal accuracy results should be investigated using DA as the index of accuracy, provided that conceptual and measurement problems discussed

Table 5
Correlates of Interpersonal Accuracy in
Previous Studies and in the Present Study

Previous Studies	Present Study ^a
Intelligence (Taft, 1950; Vernon, 1933; Wedeck, 1947)	Verbal reasoning (.31) Intelligence, high grades, and investigative interests (.26)
Dramatic and artistic interests (Estes, 1939; Taft, 1950)	
Adjustment; emotional stability (Green, 1948; Scodell and Mussen, 1953; Taft, 1950) Good impression; sense of well being; tolerance; self- control; neuroticism (-) (Vingoe and Antonoff, 1968) Sense of well being; tolerance (Hjelle, 1969)	Freedom from self-doubt and disillusionment; accept- ance of others; tendency not to worry or become stressed (.26) Self-control (.20) Tolerance (.19) Empathy (.17) Aggression (-.17)
Social detachment; independence (Adams, 1927) Achievement via conformance (-) (Edwards and McWilliams, 1974) Task versus social orientation (Taft, 1950)	Social interests; interest in observing others (-.17) Detail orientation Self-perceived detail orien- tation (.24) Heterosexuality (-.18) Exhibition (-.17)

^aVariables (composites or individual MPPB variables) are included in this part of the table only when they correlate with overall DA at the .05 level or greater.

here and elsewhere (e.g., Gage & Cronbach, 1955) can be alleviated in other person perception settings.

A possible spinoff of the research reported here is that knowledge of the person perception process (e.g., Tagiuri, 1969) may well be enriched by investigating the rating "style" or the cognitive processes relatively accurate perceivers use in making interpersonal judgments

about others. Instead of studying interpersonal perception process *rather than* accuracy, or accuracy *rather than* process, knowledge of some aspects of interpersonal perception can perhaps best be gained by studying both together.

Hopefully, the research approach used here will stimulate more ideas about how to avoid conceptual and methodological problems

related to studying interpersonal accuracy. Then more can be learned about this phenomenon, especially the judgment processes underlying the accurate perception of others.

References

- Allport, G. W. *Personality: A psychological interpretation*. New York: Henry Holt, 1937.
- Borman, W. C. Consistency of rating accuracy and rating errors in the judgment of human performance. *Organizational Behavior and Human Performance*, 1977, 20, 258–272.
- Borman, W. C. Exploring upper limits of reliability and validity in job performance ratings. *Journal of Applied Psychology*, 1978, 63, 135–144.
- Borman, W. C., Hough, L. M., & Dunnette, M. D. *Performance ratings: An investigation of reliability, accuracy, and relationships between individual differences and rater error*. Minneapolis: Personnel Decisions, Inc., 1976.
- Bruner, J. S., & Tagiuri, R. The perception of people. In G. Lindzey (Ed.), *Handbook of social psychology*. Reading, MA: Addison-Wesley, 1954.
- Cline, V. A. Interpersonal perception. In B. A. Maher (Ed.), *Progress in experimental personality research* (Vol. 1). New York: Academic Press, 1964, 221–284.
- Cronbach, L. J. Processes affecting scores on “understanding of others” and “assumed similarity.” *Psychological Bulletin*, 1955, 52, 177–193.
- Crow, W. J., & Hammond, K. R. The generality of accuracy and response sets in interpersonal perception. *Journal of Abnormal and Social Psychology*, 1957, 54, 384–390.
- Dunnette, M. D. *Forms of interpersonal accommodation: Processes, problems and research avenues*. Paper presented at the meeting of the American Psychological Association, San Francisco, September 1968.
- Gage, N. L., & Cronbach, L. J. Conceptual and methodological problems in interpersonal perception. *Psychological Review*, 1955, 62, 411–422.
- Gough, H. G. *Manual for the California Psychological Inventory*. Palo Alto, CA: Consulting Psychologists Press, 1957.
- Gough, H. G. *An exploratory cross-cultural study of interpersonal perception*. Paper presented at the meeting of the American Psychological Association, San Francisco, September 1968.
- Gough, H. G., & Heilbrun, A. L. *The Adjective Check List Manual*. Palo Alto, CA: Consulting Psychologists Press, 1965.
- Grief, E. B., & Hogan, R. The theory and measurement of empathy. *Journal of Counseling Psychology*, 1973, 20, 280–284.
- Hastorf, A. H., Schneider, D. J., & Polefka J. *Person perception*. Reading, MA: Addison-Wesley, 1970.
- Kavanagh, M. J., MacKinney, A. C., & Wolins, L. Issues in managerial performance: Multitrait-multimethod analyses of ratings. *Psychological Bulletin*, 1971, 75, 34–49.
- Kelly, G. A. *The psychology of personal constructs* (Vol. 2). New York: W. W. Norton & Co., 1955.
- Naylor, J. C., & Wherry, R. J. The use of simulated stimuli and the “JAN” technique to capture and cluster the policies of raters. *Educational and Psychological Measurement*, 1965, 25, 969–986.
- Sechrest, L., & Jackson, D. N. Social intelligence and accuracy of interpersonal predictions. *Journal of Personality*, 1961, 29, 167–182.
- Smith, P. C., & Kendall, L. M. Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 1963, 47, 149–155.
- Snyder, M. The self-monitoring of expressive behavior. *Journal of Personality and Social Psychology*, 1974, 30, 526–537.
- Taft, R. The ability to judge people. *Psychological Bulletin*, 1955, 52, 1–23.
- Tagiuri, R. Person perception. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (2nd Ed.; Vol. 3). Reading, MA: Addison-Wesley, 1969.
- Wherry, R. J. A new formula for predicting the shrinkage of the coefficient of multiple correlation. *Annals of Mathematical Statistics*, 1931, 2, 440–457.

Acknowledgments

Support for this research was provided by the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI), Contract DAHC 19-75-0006. Views expressed in this paper do not necessarily reflect those of the U.S. Army or any other agency of the U.S. Government. The author thanks M. A. Fischl of ARI for his support and advice; Janis Holtzman for collecting data; and Marvin D. Dunnette, Tom Janz, Leaetta Hough, Rodney Rosse, and Mark Davison for their aid in this study.

Author's Address

Walter C. Borman, Personnel Decisions Research Institute, 2415 Foshay Tower, Minneapolis, MN 55402.