

Empathic Accuracy

William Ickes

University of Texas, Arlington

ABSTRACT People are motivated to understand each other's psychological states as well as each other's personality traits. As a consequence, the more traditional study of accuracy in trait inference can be complemented by, and potentially benefit from, the insights provided by the more recent study of empathic accuracy. Findings in this area suggest that future research should devote more attention to (a) the history of the perceiver-target relationship; (b) the perceiver's desired future relationship with the target; (c) the possibility that perceivers have little or no "metaknowledge" regarding their own empathic ability; and (d) the possibility that, under certain conditions, perceivers might be motivated to be inaccurate, rather than accurate, in their inferences about other people's dispositions.

People are not just motivated to understand each other's personality traits; they seek to understand each other's psychological states as well. For this reason, the process of getting to know other people involves more than making correct inferences about such stable and enduring dispositions as their abilities and aptitudes, their traits and temperament, and their long-term motives and goals. It also involves making correct inferences about such unstable and transient dispositions as the thoughts they are having, the feelings they are experiencing, and the more immediate, short-term goals they are pursuing. Indeed, it may be reasonable to assume that the accurate perception of states is a necessary (if not sufficient) prerequisite to the accurate perception of traits (e.g., how can I know that you are consistently morose unless I know that you are in a bad mood today, just as you were yesterday?).

This article was written while the author was on sabbatical, and the support of the Department of Psychology, University of Washington, is gratefully acknowledged. Correspondence should be sent to William Ickes, Department of Psychology, University of Texas, Arlington, TX 76019-0528.

Journal of Personality 61:4, December 1993. Copyright © 1993 by Duke University Press. CCC 0022-3506/93/\$1.50

The Accuracy of Dispositional Judgment: Four Areas of Study

Because people are concerned about the accuracy of their state inferences as well as their trait inferences, there is more than one area of study that is relevant to the accuracy of personality judgment. Oversimplifying a bit, we can readily identify at least four of them:

1. The first area, which has the longest history of empirical study, focuses on perceivers' accuracy in judging other people's personality traits. Research in this area relies on interrater consensus as necessary but not sufficient evidence for accuracy in trait inference (e.g., Asch, 1946; Bronfenbrenner, Harding, & Gallwey, 1958; Cronbach, 1955; Estes, 1938; Funder & Colvin, 1988; McCrae, 1982; Norman & Goldberg, 1966).

2. The second area, which has a shorter and more recent history of study, focuses on dyad members' accurate perception or understanding of each other's attitudes, values, and self-conceptions. Research in this area involves comparisons of the dyad members' direct perspectives with their partners' "metaperspectives" regarding these relatively stable dispositions (e.g., Knudson, Sommers, & Golding, 1980; Laing, Phillipson, & Lee, 1966; Newmark, Woody, & Ziff, 1977; Rogers & Dymond, 1954; Sillars, 1989; Sillars & Scott, 1983).

3. The third area, which has an even more recent history of study, focuses on perceivers' accuracy or "affective sensitivity" in inferring the emotional state(s) of one or more target persons (e.g., Costanzo & Archer, 1989; Ekman & Friesen, 1975; Hall, 1978; Kagan, 1977a, 1977b; Noller, 1980, 1981; Noller & Venardos, 1986; Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979).

4. The fourth area, which is only now emerging as a field of study, focuses on perceivers' empathic accuracy—i.e., their ability to accurately infer the specific content of another person's thoughts and feelings (e.g., Ickes, Stinson, Bissonnette, & Garcia, 1990; Marangoni, Garcia, & Ickes, 1993; Simpson, Ickes, & Blackstone, 1993; Stinson & Ickes, 1992).¹

1. Because of their common focus on psychological states, the affective sensitivity research and the more recent empathic accuracy research may differ primarily in their methodology (e.g., exposure to posed and unposed facial expressions in the affective sensitivity research vs. exposure to actual social interactions in the empathic accuracy research). On the other hand, there is a definite conceptual demarcation between the current work on empathic accuracy and the earlier work on accurate empathy by Rogers

As the reader may have noticed, there is an interesting parallel between the historical order in which empirical work began in each of these areas and the type of disposition which is the focus of study in each. The first, and oldest, research area concerns perceivers' accuracy in judging those dispositions that have traditionally been viewed as among the most stable and enduring—personality traits. The second, and next-oldest, research area concerns peoples' accuracy in judging dispositions that are generally seen as somewhat less stable and enduring—attitudes, values, and self-conceptions. The third, and more recent, area concerns judgments about dispositions that are even more unstable—emotional states. And the fourth, most recent area concerns judgments about the most transient of dispositions—thoughts and feelings.

Why did researchers who were interested in accuracy in interpersonal perception begin by studying the more stable and enduring dispositions and then only gradually move toward studying the more unstable and transient ones? Most likely, the reasons are both theoretical and methodological. The theoretical reason reflects the assumption by many psychologists that perceivers might place greater reliance on the accuracy of their trait inferences than on the accuracy of their state inferences when predicting the behavior of other people. If lay perceivers—like many psychologists—believe that the more stable and enduring causes are the ones that are likely to have the greatest long-term utility in predicting behavior, then it makes sense that the study of accuracy in trait inference should have taken precedence over the study of accuracy in state inference.

Researchers should also have favored the study of trait inference over state inference on methodological grounds. First, because a target person's traits can be inferred solely on the basis of the perceiver's mental representation of the target's past behavior, it is typically not necessary to have the target person physically present or to provide the perceiver with actual records (video, audio, or written) of the target's behavior in order to conduct the research. In contrast, actual displays or re-

and his colleagues (e.g., Butler & Haigh, 1954; Rogers, 1954; Rogers & Dymond, 1954). Because Rogers and his colleagues used a Q-sort methodology to assess the convergence in the therapist's and the client's perception of the client's self-concept, their work is in the tradition of the second area of research, which focuses on a class of dispositions (attitudes, values, and self-conceptions) that are somewhat more stable and enduring than the thoughts and feelings that are the focus of the current research on empathic accuracy.

corded presentations of the target's behavior must typically be used in studies of state inference. Second, the perceiver's trait inferences can typically be assessed in a simple, straightforward way, by means of ratings on paper-and-pencil trait dimensions, whereas the perceiver's state inferences may require more complicated and open-ended forms of assessment (e.g., mood adjective checklists and thought-listing protocols). Third, because trait inferences presumably reflect the perceiver's implicit aggregation of behaviors displayed by the target person in different situations across time, such inferences might be expected to be more reliable than state inferences that presumably reflect much less implicit aggregation of the target's behavior.

There are, then, some fairly compelling reasons why researchers should have given greater precedence to the study of accuracy in trait inference than to the study of accuracy in state inference. On the other hand, the study of accuracy in trait inference provides a view of inferential accuracy that is necessarily incomplete. It is therefore important to supplement this view with insights derived from the study of accuracy in state inference.

Because of space limitations, I will not attempt to review and directly compare the findings in all four of the research areas described above. Instead, I will focus on only the fourth area—the study of empathic accuracy. This strategy is intended to minimize the risk that I will go over the same conceptual ground already covered by other contributors to this special issue. In addition, it will allow me the space to develop four specific examples of how the study of empathic accuracy can tell us more about the processes—both social and psychological—that underlie our inferences about other people's dispositions, i.e., their traits as well as their states.

In the sections to follow, I will review the concept of empathic accuracy, describe the problems involved in measuring this construct and the various solutions that researchers have proposed, and consider a number of research findings that may be relevant to the broader study of accuracy in interpersonal perception.

Empathic Accuracy

Some writers (e.g., Barrett-Lennard, 1981; Elliott et al., 1982; Goldstein & Michaels, 1985) have argued that it is important to distinguish among three components of the empathy process: empathic understanding, empathic expression, and empathic communication. The first com-

ponent—empathic understanding—involves the ability to accurately infer the thoughts and feelings of another person. The second component—empathic expression—involves the ability to express these inferred thoughts and feelings in terms that match the actual experience of the other person. The third component—empathic communication—involves the dialogic or dialectical aspect of the empathy process (e.g., the communicative relationship in which empathic understanding is developed and expressed).

When empathic accuracy is defined in strictly theoretical terms, it is clearly most synonymous with the first of these three components—empathic understanding. However, when empathic accuracy is operationally defined for the purpose of empirical study, it is typically necessary to expand its definition to include the second component—empathic expression—as well. The reason, of course, is that it is only when a perceiver's empathic understanding is expressed in some form that we are able to assess its accuracy. This point is developed in the following section, in which I propose that the most straightforward way to measure empathic accuracy is to compare the content of a target person's actual thoughts and feelings with the content of the corresponding inferred thoughts and feelings reported by the perceiver.

Measuring empathic accuracy: The Rogerian view

As Marangoni (1989) has noted, attempts to measure empathic accuracy in psychotherapy research have generally been guided by the Rogerian view that "it [empathy] involves being sensitive, *moment-to-moment*, to the changing felt meanings which flow in this other person . . ." (Rogers, 1975, p. 4, italics added). The widespread acceptance of the Rogerian view is evidenced in Truax and Carkhuff's (1967) assertion that "[v]irtually all theories of psychotherapy emphasize that for the therapist to be helpful he must be accurately empathic, be 'with' the client, be understanding, or grasp the patient's meaning" (p. 25).

The Rogerian view implies that an appropriate procedure for measuring empathic accuracy should meet at least three criteria (Marangoni, 1989). First, it should involve a temporally extended, repeated-measures assessment of the perceiver's empathic accuracy—one that can be used to track the development of empathy as an ongoing process. Second, it should allow the perceiver to generate his or her own inferences about the specific content of the target person's thoughts and feelings, rather than requiring the perceiver to choose from a set

of prefabricated response options provided by the experimenter. Third, empathic accuracy should be operationally defined by the degree to which the perceiver's inference matches—i.e., is congruent with—the target person's actual thought or feeling.

Psychotherapy research

Unfortunately, Marangoni's (1989) review led her to conclude that *none* of the procedures developed to measure empathic accuracy in psychotherapy research have met these three criteria. One common procedure uses as its stimulus materials audio- or videotapes of actual counseling sessions, or typed transcripts of actual or simulated psychotherapy sessions. (In some of these studies, the client's portion of the interaction is edited out, so that only the therapist's portion remains.) This procedure requires that persons who are presumed to be expert judges (e.g., clinical supervisors/trainers) each make a single global rating of "the level of empathic responsivity exhibited by student trainees" (Marangoni, 1989, p. 16). Studies using this procedure have been reviewed by Truax and Carkhuff (1967) and by Carkhuff (1969a, 1969b).

Procedures of this type do not satisfy the first criterion noted above, and it is debatable how well they satisfy either of the remaining two criteria. In addition, research using such procedures has been plagued with a number of other problems (Marangoni, 1989, pp. 23–28). First, the observer-judges who were asked to rate the therapist's empathic accuracy were, in many cases, undergraduate or graduate students in psychology whose training was either minimal or unspecified. Second, in a review of 53 studies published between 1980 and 1984, Wilson, Griswold, and Sunderland (1984) concluded that the reliability of the observers was often "either not discussed or dismissed with a claim that the observers were 'expert' and therefore their ratings were valid" (p. 42). Third, "even under the best of rater-training conditions, the accuracy criterion is still an external one, residing in the consensual agreement of outside observers" (Marangoni, 1989, p. 26). Fourth, there is a "lack of research evidence documenting the objective veracity of [the ratings made by such observers]" (Marangoni et al., 1993, p. 4) (for similar criticisms, see Carkhuff & Burstein, 1970; Kurtz & Grummon, 1972; Truax & Carkhuff, 1967; Wilson & Griswold, 1985).

Closer in its conception to the Rogerian view of accurate empathy is the Affective Sensitivity Scale developed by Kagan and his colleagues (Campbell, Kagan, & Krathwohl, 1971; Danish & Kagan, 1971; Kagan,

1972, 1977a, 1977b). The Affective Sensitivity Scale relies on a video presentation of a series of 41 brief excerpts from videotaped counseling sessions. Collectively, these excerpts portray 11 different male and female client-counselor dyads discussing a range of client problems. Subjects are asked to view these excerpts in succession and, following each, to "choose from a set of 3 options the sentence they think most accurately reflects the client's feelings toward the content of what he/she is communicating" (Marangoni, 1989, p. 29). The correct option in each set of three is a statement of the client's actual reported feeling, as assessed when the client subsequently viewed the videotaped therapy session as part of a stimulated recall procedure.

The Affective Sensitivity Scale has acceptable internal consistency and test-retest reliability (typical coefficients are in the .70s; Kagan, 1977a). In addition, this scale is the only measure of empathy developed in psychotherapy research "that uses as its accuracy criterion data obtained from the actual target for whom inferences are being generated" (Marangoni, 1989, p. 30). On the other hand, while it satisfies the third criterion implied by a Rogerian view (operationally defining empathic accuracy by the degree to which the perceiver's inference matches the target person's actual feeling), it does not satisfy either of the first two criteria. That is, it does not provide a temporally extended, repeated measure that can be used to track empathy as an ongoing process, nor does it allow the perceiver to generate his or her own inferences about the content of the target person's thoughts and feelings.²

Laboratory research

A procedure for measuring empathic accuracy that satisfies all three criteria was recently developed by the author and his colleagues (Ickes, Bissonnette, Garcia, & Stinson, 1990; Ickes & Tooke, 1988). Although its original application was in laboratory studies of unstructured dyadic interaction (Ickes, Stinson, Bissonnette, & Garcia, 1990; Simpson et al., 1993; Stinson & Ickes, 1992), it was subsequently adapted for use in a clinically relevant setting by Marangoni et al. (1993).

2. Analogous to the manner in which a reconstruction task is a more stringent test of memory than a recognition task, allowing the perceiver to generate his or her own inferences about a target's thoughts and feelings is a more stringent test of empathic accuracy than allowing the perceiver to choose from a set of prefabricated response options.

The procedure's original application, in the context of the dyadic interaction paradigm (Ickes, Bissonnette, Garcia, & Stinson, 1990), can be described as follows. The members of each dyad—whether strangers or partners in an established relationship—are led into a waiting room and are left there together in the experimenter's absence. During this interval (in which the subjects are ostensibly waiting for the experiment to begin), the subjects' verbal and nonverbal behaviors are unobtrusively audio- and videotaped. When the experimenter returns at the end of the observation period, the subjects are partially debriefed and asked for their signed consent to release the tape for use as data. They are also asked to participate in a second part of the study that concerns their specific thoughts and feelings during the interaction.

If their signed consent is given, the subjects are then seated in separate but identical cubicles where they are each instructed to view a separate videotape of the interaction. By stopping the videotape with a remote start/pause control at those points where they remember having had a specific thought or feeling, each subject makes a written, time-logged listing of these *actual thought/feeling entries*. The subjects are then instructed to view the videotape a second time, and the tape is stopped for them at each of those points at which their interaction partner had reported a thought or feeling. The subject's task during this phase is to infer the content of their partner's thoughts and feelings and provide a written, time-logged listing of these *inferred thought/feeling entries*. When both subjects have completed this task and have filled out a short post-test questionnaire, they are debriefed more completely and then thanked and released.

The data collected during this procedure are subsequently used to compute a measure of empathic accuracy. The computation of this measure requires similarity judgments to be made by trained, independent raters. With the aid of a custom software program created by Victor Bissonnette, all actual thought/feeling entries and the corresponding inferred thought/feeling entries are presented as paired stimuli on the screen of a microcomputer. The task of the independent raters is to judge the similarity of each pair on a 3-point scale ranging from 0 (essentially different content) through 1 (similar but not the same content) to 2 (essentially the same content). For examples of empathic accuracy judgments at all three levels of similarity, see Table 1.

In studies using six independent raters to make these similarity judgments, the internal consistency of the raters' judgments was very high (.94 in the study by Ickes, Stinson, Bissonnette, & Garcia, 1990, and

Table 1
Sample Thought/Feeling Entries with Corresponding Inferences
and Mean Empathic Accuracy Ratings

Dyad member's actual thought or feeling	Partner's inference	Mean empathic accuracy rating (minimum = 0, maximum = 2)
I was thinking about a previous production of the play in another city that a local radio personality was in.	She was thinking if I would ask her out.	0
I was thinking that I was not missing anything I didn't want to miss. I was thinking that I came to school to learn, not to join organizations.	He was thinking about what he was missing at school.	1
I was feeling silly because I couldn't remember my instructor's name.	She was maybe feeling sorta odd for not remembering her teacher's name.	2

.95 in the study by Stinson & Ickes, 1992). In a study using only four independent raters to make these judgments, the internal consistency was only slightly lower at .85 (Marangoni et al., 1993).³ These data indicate that empathic accuracy can be measured very reliably with the procedure just described. They also suggest that the raters in these studies use essentially the same criteria when judging the similarity between actual versus inferred thought/feeling entries.

Once the internal consistency data have established that these similarity judgments are reliable, the data are then averaged across raters in order to compute mean content accuracy scores for each of the actual and inferred thought/feeling pairs. If a global (i.e., aggregated) measure of empathic accuracy is desired, these mean ratings are first summed across all of the thought/feeling inferences in a given subject's protocol. These summed values are then divided by the maximum num-

3. Raters, rather than entries, were treated as items in the calculation of Cronbach's alpha in these studies.

ber of accuracy points that could be obtained for a given number of inferences in order to derive an overall accuracy score that controls for individual differences in the number of inferences made. This global accuracy score is conveniently scaled, with a possible range of .00 (total inaccuracy) to 1.00 (total accuracy).

This procedure for assessing empathic accuracy satisfies all three of the criteria implicit in a Rogerian view of empathy (Marangoni, 1989). First, it permits a temporally extended, repeated-measures assessment of the perceiver's empathic accuracy—one that can be used to track the development of empathy as an ongoing process. Second, it allows the perceiver to generate his or her own inferences about the specific content of the target person's thoughts and feelings, rather than requiring the perceiver to choose from a set of prefabricated response options provided by the experimenter. Third, it operationally defines empathic accuracy as the degree to which the perceiver's inferences match—i.e., are congruent with—the target person's actual thoughts and feelings.

To date, this procedure has been used in four studies. Three of these are studies of unstructured dyadic interaction; the fourth is a clinical analog study in which our procedure for assessing empathic accuracy was adapted by Marangoni et al. (1993) for use in a clinically relevant setting.

Ickes, Stinson, Bissonnette, and Garcia (1990). In the first dyadic interaction study, Ickes, Stinson, Bissonnette, and Garcia (1990) examined correlates of empathic accuracy in the initial, unstructured interactions of 38 mixed-sex (male-female) dyads. The results revealed that, for these strangers of the opposite sex, empathic accuracy was positively correlated with (a) the amount of behavioral involvement (e.g., talking and directed gaze) displayed by the dyad members; (b) the grade point average of the perceiver; (c) the physical attractiveness of the target person (i.e., the partner); and (d) the percentage of partner attributions and partner-relevant thoughts and feelings which the perceiver reported.

Stinson and Ickes (1992). In the second dyadic interaction study, Stinson and Ickes (1992) found that male friends were more accurate than male strangers in inferring each other's thoughts and feelings. Plausible reasons for this difference were that the friends interacted more and exchanged more information; had more similar personalities and therefore had more rapport with each other; and had more extensive and detailed knowledge about the structure and contents of each other's memory.

The results confirmed that the male friends did indeed interact more and were more similar in their self-reported sociability than the male strangers; however, these differences did not account for the friends' greater empathic accuracy. Instead, the friends' advantage in this regard appeared to be attributable to the greater knowledge they had about the structure and contents of each other's memory (cf. Wegner, 1987; Wegner, Giuliano, & Hertel, 1985).

Simpson, Ickes, and Blackstone (1993). In the third dyadic interaction study, Simpson et al. (1993) tested the empathic accuracy of 82 dating couples during a laboratory task that was systematically varied to create different levels of perceived threat to their relationships. In both conditions of the study, the members of these dating couples took turns viewing and rating slides of target individuals of the opposite sex. These target individuals were described as undergraduates at the subjects' university who were currently available as dating partners, and who had agreed not only to let their photographs be used in the research but also to be interviewed by some of the subjects at a later time.

In the high-threat condition, the opposite-sex targets were extremely attractive in both sets of slides, i.e., those rated by the male and those rated by the female dating partners. In the low-threat condition, on the other hand, the targets were all below average in their attractiveness. The subjects rated the targets aloud (on the dimensions of physical attractiveness and sexual appeal) in the presence of their dating partners. The subjects' own thoughts and feelings and their inferences about their partners' thoughts and feelings were subsequently assessed using the same type of video-cued recall procedure described earlier.

The results revealed that empathic accuracy was lower for couples in the high-threat condition than for those in the low-threat condition. Empathic accuracy was also lower for couples who described their relationship as high versus low in closeness, and as insecure rather than secure. These effects were additive such that empathic accuracy was significantly lower for couples in the high-threat, high-close, high-insecure condition than for couples in the remaining conditions combined. These findings appear to be consistent with a motivated inaccuracy interpretation, which suggests that when their dating relationship is perceived as close but insecure and situationally threatened, individuals may attempt to preserve the relationship by denying or otherwise failing to acknowledge their partners' feelings of attraction to potential alternative partners. Moreover, ancillary analyses ruled out the two most plausible

alternative explanations (i.e., one positing greater concealment of the partner's actual thoughts and feelings, and one positing greater stress which disrupts the empathic accuracy of the perceiver).

Marangoni, Garcia, and Ickes (1993). After our procedure for assessing empathic accuracy had been used in the first two dyadic interaction studies, it was adapted for use in a clinically relevant setting by Marangoni et al. (1993). In this study, 80 undergraduates viewed a series of three simulated psychotherapy sessions. In each session, a different female client discussed her real-life problem(s) with a professional client-centered therapist. (Immediately after each session had been taped, the client had reviewed the videotape in a cued-recall procedure and had made a complete, time-logged listing of all of the actual thoughts and feelings that she had experienced during the session.)

As the subjects viewed them, the tapes were put on pause at each of those points at which the client had reported having had a specific thought or feeling. The subjects' task was to infer the content of the thought or feeling and to record it in writing on a thought/feeling reporting form. Half of the subjects received no feedback about the accuracy of the inferences they made (no feedback condition), whereas the other half received feedback during the middle portion of each tape (feedback condition). For the subjects in the feedback condition, a sentence reporting the client's actual thought or feeling appeared on the television monitor following each of the subject's inferences.

This study addressed three questions regarding empathic accuracy in a clinically relevant setting. First, does the empathic accuracy of a perceiver improve with increasing exposure to a target individual? Second, can empathic accuracy be enhanced by providing the perceiver with feedback about the target's actual thoughts and feelings? Third, are there stable individual differences in empathic accuracy that generalize across different targets? The results indicated that although absolute performance levels varied from one target (i.e., client) to the next, empathic accuracy generally improved with increasing exposure to the target. In addition, feedback about the target's actual thoughts and feelings accelerated the rate at which the perceivers' empathic accuracy improved. Finally, cross-target consistency in responding ($\alpha = .86$) revealed stable individual differences in the perceivers' empathic ability.

Implications for the Broader Study of Inferential Accuracy

I believe that the study of empathic accuracy offers a number of useful insights from which the broader study of inferential accuracy can benefit. By way of illustration, I would like to examine four implications that are suggested by findings of the studies I have just reviewed. The first implication is that the history of the relationship between the perceiver and the target is extremely important, and deserves to be taken more seriously than much of the previous work on dispositional inference would suggest. A second implication is that the perceiver's desired future relationship with the target can also exert important, though seldom recognized, influences on the perceiver's inferences about the target. A third implication is that the inferential ability of the perceiver may be just as important as previous theorists and researchers have assumed, but that its study may be complicated by the failure of perceivers to provide accurate self-reports regarding this ability (i.e., by their lack of "metaknowledge"). A fourth implication is that the study of dispositional inference may be further complicated by instances of motivated inaccuracy on the part of the perceiver. Let's consider each of these implications in turn.

History of the perceiver-target relationship

The importance of the history of the relationship between the perceiver and the target is indicated by a number of findings that have emerged in our studies of empathic accuracy. For example, Stinson and Ickes (1992) found that, on average, male friends had empathic accuracy scores that were about 50% higher than those of male strangers ($M_s = 29.8$ vs. 19.9 , $p < .02$). This finding may not be surprising; however, it does raise the interesting question of *why* friends should be more accurate than strangers in inferring each other's thoughts and feelings.

One hypothesis considered by Stinson and Ickes (1992) is that friends have more similar personalities than strangers, and therefore experience a *similarity-based rapport* that enhances their empathic accuracy. This hypothesis seemed plausible because the male friends in the Stinson and Ickes (1992) study were found to be more similar in their sociability than the male strangers (i.e., the intraclass correlation of the friends' sociability scores, .65, was significantly greater than that of the

strangers' scores, $-.08$).⁴ The appropriate covariance analysis did not support this hypothesis, however. Statistically controlling for the effect of this personality similarity did not substantially attenuate the correlation between relationship status (friends vs. strangers) and empathic accuracy.

A second hypothesis is that friends exchange more information than strangers do during their immediate, unstructured interactions, and that this *immediate information* accounts for their greater empathic accuracy. This hypothesis seemed plausible because the male friends in the Stinson and Ickes (1992) study displayed considerably more interactional involvement than the male strangers (i.e., they talked, looked, smiled, and gestured significantly more). A covariance analysis again revealed, however, that statistically controlling for the effect of interactional involvement did not substantially attenuate the correlation between relationship status (friends vs. strangers) and empathic accuracy. Moreover, while the empathic accuracy of the male strangers was strongly dependent on the level of interactional involvement in their immediate interaction, $r = .50$, the empathic accuracy of the male friends was not, $r = .09$. This pattern of findings suggests that the greater empathic accuracy of the male friends derives from an understanding of each other that transcends the information available in a single, specific interaction.

Thus, a third hypothesis is that friends have more detailed and extensive knowledge of each other's lives than do strangers, and that this difference in *knowledge structures* mediates the observed difference in empathic accuracy. This hypothesis suggested that the limitations of strangers' knowledge structures should be particularly evident when they were asked to make inferences about their partners' thoughts and feelings about imagined events occurring in another place or another time (AP-AT). Specifically, the greater the percentage of AP-AT entries reported by their partners, the worse the strangers' empathic accuracy scores should be; thus, a significant negative correlation was expected. In contrast, because the interaction-based cues that evoke AP-AT thoughts and feelings in one friend should tend to evoke the same thoughts and feelings in the other, the percentage of AP-AT entries re-

4. Sociability was the only 1 out of 14 personality dimensions tested by Stinson and Ickes (1992) for which the intraclass correlation of the friend's scores was significantly greater than that of the stranger's scores.

ported by their partners was expected to be positively correlated with the friends' empathic accuracy scores.

These predictions were confirmed. The percentage of AP-AT thoughts and feelings reported by their partners was negatively correlated with the strangers' empathic accuracy, $r = -.57$, but positively correlated with the friends' empathic accuracy, $r = .33$, and the difference between these correlations was clearly significant, $p < .005$. These findings, along with those from relevant follow-up analyses, indicate that the greater empathic accuracy of the friends was in large measure accounted for by the friends' ability to accurately read their partners' thoughts and feelings about imagined events in another place or time, and by the strangers' inability to do the same. A similar finding was obtained by Colvin and Funder (1991), who reported that strangers were as able as friends to predict behavior in a future situation similar to the one in which strangers (but not friends) had observed the targets, but that friends were much better at predicting behavior more generally across settings that were *not* similar to those the strangers observed.

The importance of the history of the relationship between the perceiver and the target was further indicated in the Stinson and Ickes (1992) study by the finding that the empathic accuracy scores of the male friends were clearly interdependent (intraclass $r = .37$). This finding suggests that the friends had developed converging levels of understanding of each other. In contrast, the empathic accuracy scores of the male strangers were independent (intraclass $r = -.04$), revealing no evidence of convergence in their levels of understanding.

Interestingly, a similar finding emerged in Simpson et al.'s (1993) study of 82 dating couples. For stable couples (those in which both members reported that they were still dating 3 months later), the partners' empathic accuracy scores were significantly correlated, $r = .29$. However, for unstable couples (those in which both members reported that they were no longer dating 3 months later), the partners' empathic accuracy scores were not correlated, $r = -.10$. This finding suggests that stable dating couples, like close, same-sex friends, achieve similar levels of understanding that unstable couples, like same-sex strangers, do not achieve. It remains to be determined, however, whether this convergence in understanding is a cause or a consequence of relationship stability, or whether these two variables exert a reciprocal influence on each other. It also remains to be determined whether such convergence in the partners' understanding can exist from the very beginning

of their relationship, or whether it must develop over time. And, if the convergence develops over time, does the less-accurate partner's understanding always increase to match the level of the more-accurate partner, or can convergence in the opposite direction also occur?⁵

In summary, the history of the perceiver-target relationship has important implications for the degree to which the participants understand each other, and the degree to which their levels of understanding are interdependent rather than independent. These implications may apply to other forms of dispositional inference as well. Friends, for example, may not only make more valid and differentiated trait attributions about each other than strangers do, but may also exhibit more convergence in their views of self and other. This intersubjective aspect of dispositional inference should, in the future, be a topic of interest to all researchers who study accuracy in interpersonal perception.

The desired future relationship with the target

Findings from the studies by Ickes, Stinson, Bissonnette, & Garcia (1990) and Simpson et al. (1993) suggest that a perceiver's desired future relationship with the target can also exert important influences on empathic accuracy. In the Ickes et al. (1990) study, it was found that the subjects' accuracy in inferring the thoughts and feelings of the opposite-sex strangers with whom they were paired was positively correlated with the physical attractiveness of these partners, $r = .24$, $p < .05$. This finding suggests that the desire to establish a positive relationship with the target might provide one motive for attempting to accurately infer the target's thoughts and feelings.⁶

5. It is possible that relationships are inherently more unstable between partners who differ greatly in their empathic accuracy than between partners whose levels of empathic accuracy are more convergent. Inequities in the feeling of being understood by the partner, etc., may be more likely in the first type of relationship than in the second. Disparities in empathic accuracy might also reflect flawed intersubjectivity in relationships, a situation in which partners cannot bring their individual understandings into synchrony with each other.

6. The correlation between partner attractiveness and empathic accuracy is especially noteworthy when one considers that empathic accuracy was assessed after the interaction by means of a procedure designed to ensure that all subjects would be highly motivated to accurately infer their interaction partner's thoughts and feelings. Apparently, the subjects' attraction to their opposite-sex partners affected their attentiveness *during* the interaction in a way that heightened their empathic accuracy during the assessment that occurred later on.

What happens, however, when an accurate assessment of the other person's thoughts and feelings is potentially threatening to a positive relationship that has already been established? Simpson et al. (1993) addressed this question in their study of 82 dating couples. They found that empathic accuracy was especially low for couples who perceived their dating relationships to be close but insecure and highly threatened by their partner's possible attraction to the physically and sexually appealing target persons whom they were evaluating as potential dates. This finding suggests that the desire to maintain a good future relationship might motivate individuals to be less than accurate in acknowledging their partners' feelings of attraction to potential alternative partners.

Apparently, the perceiver's desire for a good future relationship with the target can in some conditions enhance, but in other conditions impair, the perceiver's accuracy in inferring the target's thoughts and feelings. Similar effects might also be found in studies of more stable dispositional inferences. For example, consistent with Dale Carnegie's (1936) advice, perceivers might take special care to accurately note and comment on the traits and attributes of targets with whom they desire a closer relationship. Once such a relationship has been established, however, perceivers may be reluctant to accurately perceive some of their partners' more negative traits and attributes—particularly those which, if consciously acknowledged, might threaten the established relationship.

Lack of metaknowledge regarding empathic accuracy

In general, the evidence from the studies my colleagues and I have conducted suggests that people lack metaknowledge regarding their own empathic accuracy. For example, in the study by Ickes, Stinson, Bissonnette, & Garcia (1990), various self-report measures of empathic skills and empathic accuracy proved to be disappointingly poor predictors of actual empathic accuracy. The relevant correlations, reported in Table 2, indicate that neither the subscales of Davis's (1983) empathy measure nor the empathic accuracy measure developed by Ickes (1988) were significant predictors of empathic accuracy for either the male or the female subjects in this study. Worse yet, most of these correlations were negative rather than positive.

The clinically relevant study by Marangoni et al. (1993) offered further evidence that people may be unreliable judges of their own empathic ability. At the end of the session, after the subject had attempted

Table 2
Correlations between Self-Reported Empathic Skills and Dispositions and Actual Empathic Accuracy in the Ickes, Stinson, Bissonnette, and Garcia (1990) Study

Self-report measures	Actual empathic accuracy		
	Men	Women	Combined
Perspective taking (PT)	-.15	-.15	-.14
Empathic concern (EC)	.09	-.09	.04
Fantasy identification (F)	-.13	-.16	-.11
Personal distress (PD)	.16	-.12	.06
Empathic accuracy (EA)	.10	-.21	-.03

Note. None of the correlations reported in this table was statistically significant. Adapted by permission of the publisher from "Naturalistic social cognition: Empathic accuracy in mixed-sex dyads" by W. Ickes, L. Stinson, V. Bissonnette, and S. Garcia, 1990, *Journal of Personality and Social Psychology*, 59, pp. 730-742. Copyright 1990 by the American Psychological Association.

Table 3
Correlations between Self-Rated Empathic Accuracy and Actual Empathic Accuracy for Each of the Clients Viewed in the Marangoni et al. (1993) Study

Self-rated accuracy	Actual empathic accuracy		
	Men	Women	Combined
Client 1	.08	.07	.08
Client 2	-.05	.04	.00
Client 3	-.22	-.10	-.18

Note. None of the correlations reported in this table was statistically significant (adapted from Marangoni, Garcia, and Ickes [1993]).

to infer the thoughts and feelings of each of the three female clients, the subject was asked to provide scaled ratings of his or her empathic accuracy. A brief videoclip of each female client was presented, following which the subject was asked to respond, on a 5-point scale, to the question "How well do you think you inferred [that particular client's] thoughts and feelings?"

Once again, the results revealed little or no relationship between self-rated empathic accuracy and actual empathic accuracy (see Table 3). These null findings are striking given that (a) the subjects were asked

to estimate their empathic accuracy with respect to particular targets, instead of targets in general; and (b) the subjects displayed reliable individual differences in their actual empathic accuracy—differences that were highly stable ($\alpha = .86$) across the three female clients. Collectively, these findings suggest that although people differ reliably in their ability to accurately read the thoughts and feelings of others, they may have little insight regarding their own relative level of empathic skill.

Reasons for this lack of metaknowledge can be readily suggested. A few possibilities are that (a) to avoid violating privacy norms, perceivers rarely seek explicit feedback about their accuracy in inferring other people's thoughts and feelings; (b) such feedback, when it is provided, rarely concerns that perceiver's *relative* level of empathic accuracy (i.e., the perceiver's rank order in comparison to other perceivers); (c) the verbal and nonverbal feedback that perceivers *do* receive from targets may be misleading (i.e., to be polite or avoid giving offense, targets may provide head nods and other signs of agreement or acknowledgment even when the perceiver has failed to understand them); (d) targets may also mask—or simply fail to display—the cues that signal that they are currently experiencing a covert thought or feeling; (e) perceivers may eventually achieve some insight regarding the degree to which they understand their few closest friends and intimates, but mistakenly believe that this level of ability generalizes to other targets as well; and (f) to the degree that perceivers are egocentric, both their empathic accuracy and their metaknowledge of this ability may be impaired.

Ironically, then, while the inferential ability of the perceiver may be just as important as previous theorists and researchers have assumed, its study may be complicated by the failure of perceivers to provide accurate self-reports regarding this ability. This same lack of metaknowledge may help to account for the general lack of behavioral validity evidence for the various self-report measures of empathy (Hogan, 1969; Kerr & Speroff, 1954), empathic ability (Davis, 1983; Dymond, 1949), empathic accuracy (Ickes, 1988), social intelligence (Walker & Foley, 1973), and social insight (Chapin, 1942) that researchers have developed during the past 50 years. In other words, if self-report measures consistently fail to predict inferential accuracy, as reviewers have repeatedly noted (e.g., Chlopan, McCain, Carbonell, & Hagen, 1985; Funder & Harris, 1986; Goldstein & Michaels, 1985; Kenny & Albright, 1987; Marangoni, 1989; Walker & Foley, 1973), it may be time to acknowledge the possibility that the fault lies not in the methods used to construct these self-report measures but in the

assumption that subjects have any real insight regarding this ability. A better approach might be to measure inferential accuracy directly and to assess its generality across different targets, as Marangoni et al. (1993) have done.

In addition, it may be worthwhile to reexamine the literature on the accuracy of inferences about more stable dispositions such as traits, attitudes, and values. If the evidence suggests that (a) people differ reliably in their ability to accurately infer these more stable dispositions, but (b) efforts to find self-report correlates of such accuracy have generally failed, then the argument for deficient metaknowledge becomes even stronger.

Motivated inaccuracy on the part of the perceiver

The findings from Simpson et al.'s (1993) study of 82 dating couples raise the intriguing possibility that, under certain conditions, perceivers might be motivated to be inaccurate, rather than accurate, in their inferences about a target person's thoughts and feelings. If cases of motivated inaccuracy can be identified and substantiated in the research on state inference, it may be profitable to look for them in the research on trait inference as well. For example, if close but insecure dating partners are unwilling to perceive each other's feelings of attraction to others and potential for sexual disloyalty, it may also be the case, for example, that hospital staff would be reluctant to perceive a homicidal disposition in a nurse who has been entrusted with the care of highly vulnerable patients, or that bank managers would be reluctant to perceive a larcenous disposition in a long-term employee who is in a position to embezzle.

Assuming that cases of motivated inaccuracy *can* be identified and substantiated, theorists will still have to grapple with the thorny conceptual problem of how perceivers are able to keep themselves from consciously knowing things that they would prefer not to know. That problem is clearly beyond the scope of the present article, but it represents a challenging issue for future theory and research.

CONCLUSION

The study of accuracy in dispositional inference involves more than the study of how people infer each other's personality traits. It also implicates the study of how people infer each other's psychological states,

i.e., their thoughts and their feelings. Indeed, the perception of consistency in the psychological states of other people may be an important basis of accurate trait inference. For this reason, the more traditional study of accuracy in trait inference can be complemented by, and potentially benefit from, the more recent study of empathic accuracy. Work in this area has already made salient a number of issues that may have implications for the broader study of inferential accuracy. These issues highlight the importance of (a) the history of the perceiver-target relationship, (b) the perceiver's desired future relationship with the target, (c) the perceiver's lack of metaknowledge regarding her or his own empathic ability, and (d) the perceiver's motivation, in certain conditions, to make inaccurate rather than accurate inferences about another person's dispositions.

REFERENCES

- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, **41**, 258-290.
- Barrett-Lennard, G. T. (1981). The empathy cycle: Refinement of a nuclear concept. *Journal of Counseling Psychology*, **39**, 333-339.
- Bronfenbrenner, U., Harding, J., & Gallwey, M. (1958). The measurement of skill in social perception. In D. McClelland (Ed.), *Talent and society: New perspectives in the identification of talent*. New York: Van Nostrand.
- Butler, J. M., & Haigh, G. V. (1954). Changes in the relationship between self-concepts and ideal concepts consequent upon client-centered counseling. In C. R. Rogers & R. F. Dymond (Eds.), *Psychotherapy and personality change*. Chicago: University of Chicago Press.
- Campbell, R. J., Kagan, N., & Krathwohl, D. R. (1971). The development and validation of a scale to measure affective sensitivity (empathy). *Journal of Counseling Psychology*, **18**, 407-412.
- Carkhuff, R. R. (1969a). *Helping and human relations: A primer for lay and professional helpers: Vol. 1. Selection and training*. New York: Holt, Rinehart & Winston.
- Carkhuff, R. R. (1969b). *Helping and human relations: A primer for lay and professional helpers: Vol. 2. Practice and research*. New York: Holt, Rinehart & Winston.
- Carkhuff, R. R., & Burstein, J. (1970). Objective therapist and client ratings of therapist-offered facilitative conditions of moderate to low functioning therapists. *Journal of Clinical Psychology*, **26**, 394-395.
- Carnegie, D. (1936). *How to win friends and influence people*. New York: Simon & Schuster.
- Chapin, F. S. (1942). Preliminary standardization of a social insight scale. *American Sociological Review*, **7**, 214-225.
- Chlopan, B. E., McCain, M. L., Carbonell, J. L., & Hagen, R. L. (1985). Empathy: Review of available measures. *Journal of Personality and Social Psychology*, **48**, 635-653.

- Colvin, C. R., & Funder, D. C. (1991). Predicting personality and behavior: A boundary on the acquaintanceship effect. *Journal of Personality and Social Psychology*, **60**, 884-894.
- Costanzo, M., & Archer, D. (1989). Interpreting the expressive behavior of others: The Interpersonal Perception Task. *Journal of Nonverbal Behavior*, **13**, 225-245.
- Cronbach, L. J. (1955). Processes affecting scores on "understanding of others" and "assumed similarity." *Psychological Bulletin*, **52**, 177-193.
- Danish, S. J., & Kagan, N. (1971). Measurement of affective sensitivity: Toward a valid measure of interpersonal perception. *Journal of Counseling Psychology*, **18**, 51-54.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, **44**, 113-126.
- Dymond, R. F. (1949). A scale for the measurement of empathic ability. *Journal of Consulting Psychology*, **13**, 127-133.
- Ekman, P., & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial cues*. Englewood Cliffs, NJ: Prentice-Hall.
- Elliott, R., Filipovich, H., Harrigan, L., Gaynor, J., Reimschuessel, C., & Zapadka, J. K. (1982). Measuring response empathy: The development of a multi-component rating scale. *Journal of Counseling Psychology*, **29**, 379-387.
- Estes, S. G. (1938). Judging personality from expressive behavior. *Journal of Abnormal and Social Psychology*, **33**, 217-236.
- Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology*, **55**, 149-158.
- Funder, D. C., & Harris, M. J. (1986). On the several facets of personality assessment: The case of social acuity. *Journal of Personality*, **54**, 528-550.
- Goldstein, A. P., & Michaels, G. Y. (1985). *Empathy: Development, training and consequences*. Hillsdale, NJ: Lawrence Erlbaum.
- Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, **85**, 845-857.
- Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, **33**, 307-316.
- Ickes, W. (1988). *An 18-item measure of empathic accuracy*. Unpublished scale and associated psychometric data, University of Texas, Arlington.
- Ickes, W., Bissonnette, V., Garcia, S., & Stinson, L. (1990). Implementing and using the dyadic interaction paradigm. In C. Hendrick & M. Clark (Eds.), *Review of personality and social psychology: Research methods in personality and social psychology* (Vol. 11, pp. 16-44). Newbury Park, CA: Sage.
- Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, **59**, 730-742.
- Ickes, W., & Tooke, W. (1988). The observational method: Studying the interaction of minds and bodies. In S. Duck, D. Hay, S. Hobfoll, W. Ickes, & B. Montgomery (Eds.), *The handbook of personal relationships: Theory, research, and interventions* (pp. 79-97). Chichester, England: John Wiley.
- Kagan, N. (1972). *Influencing human interaction*. East Lansing: Michigan State University Press.

- Kagan, N. (1977a). *Affective sensitivity test: Validity and reliability*. Paper presented at the convention of the American Psychological Association, San Francisco.
- Kagan, N. (1977b). *Interpersonal process recall*. East Lansing: Michigan State University Press.
- Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin*, **102**, 390-402.
- Kerr, W. A., & Speroff, B. J. (1954). Validation and evaluation of the empathy test. *Journal of General Psychology*, **50**, 269-276.
- Knudson, R. A., Sommers, A. A., & Golding, S. L. (1980). Interpersonal perception and mode of resolution in marital conflict. *Journal of Personality and Social Psychology*, **38**, 251-263.
- Kurtz, R. R., & Grummon, D. L. (1972). Different approaches to the measurement of therapist empathy and their relationship to therapy outcomes. *Journal of Consulting and Clinical Psychology*, **39**, 106-115.
- Laing, R. D., Phillipson, H., & Lee, A. R. (1966). *Interpersonal perception: A theory and a method of research*. New York: Springer.
- Marangoni, C. (1989). *The facilitative effect of feedback training on objectively assessed empathic accuracy in a clinically relevant context*. Unpublished doctoral dissertation, University of Texas, Arlington.
- Marangoni, C., Garcia, S., & Ickes, W. (1993). *Empathic accuracy in a clinically relevant setting*. Manuscript submitted for publication.
- McCrae, R. R. (1982). Consensual validation of personality traits: Evidence from self-reports and ratings. *Journal of Personality and Social Psychology*, **43**, 293-303.
- Newmark, C. S., Woody, G., & Ziff, D. (1977). Understanding and similarity in relation to marital satisfaction. *Journal of Clinical Psychology*, **33**, 83-86.
- Noller, P. (1980). Misunderstandings in marital communication: A study of couples' nonverbal communication. *Journal of Personality and Social Psychology*, **39**, 1135-1148.
- Noller, P. (1981). Gender and marital adjustment level differences in decoding messages from spouses and strangers. *Journal of Personality and Social Psychology*, **41**, 272-278.
- Noller, P., & Venardos, C. (1986). Communication awareness in married couples. *Journal of Social and Personal Relationships*, **3**, 31-42.
- Norman, W. T., & Goldberg, L. R. (1966). Raters, ratees, and randomness in personality. *Journal of Personality*, **4**, 681-691.
- Rogers, C. R. (1954). The case of Mrs. Oak: A research analysis. In C. R. Rogers & R. F. Dymond (Eds.), *Psychotherapy and personality change*. Chicago: University of Chicago Press.
- Rogers, C. R. (1975). Empathic: An unappreciated way of being. *Counseling Psychologist*, **5**, 2-10.
- Rogers, C. R., & Dymond, R. F. (1954). *Psychotherapy and personality change*. Chicago: University of Chicago Press.
- Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, P. L., & Archer, D. (1979). *Sensitivity to nonverbal communication: The PONS test*. Baltimore: Johns Hopkins University Press.
- Sillars, A. L. (1989). Communication, uncertainty and understanding in marriage. In B. Dervin, L. Grossberg, B. O'Keefe, & E. Wartella (Eds.), *Rethinking communication: Vol. 2. Paradigm exemplars* (pp. 307-328). Newbury Park, CA: Sage.

- Sillars, A. L., & Scott, M. D. (1983). Interpersonal perception between intimates: An integrative review. *Human Communication Research*, **10**, 153-176.
- Simpson, J., Ickes, W., & Blackstone, T. (1993). *When the head protects the heart: Motivated inaccuracy in dating relationships*. Manuscript submitted for publication.
- Stinson, L., & Ickes, W. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of Personality and Social Psychology*, **62**, 787-797.
- Taft, R. (1955). The ability to judge people. *Psychological Bulletin*, **52**, 1-23.
- Truax, C. B., & Carkhuff, R. R. (1967). *Toward effective counseling and psychotherapy: Training and practice*. Chicago: Aldine.
- Walker, R. E., & Foley, J. M. (1973). Social intelligence: Its history and measurement. *Psychological Reports*, **33**, 839-864.
- Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185-208). New York: Springer-Verlag.
- Wegner, D. M., Giuliano, T., & Hertel, P. (1985). Cognitive interdependence in close relationships. In W. Ickes (Ed.), *Compatible and incompatible relationships* (pp. 253-276). New York: Springer-Verlag.
- Wilson, F. R., & Griswold, M. L. (1985). The effects of method and comprehensiveness of training on the reliability and validity of ratings of counselor empathy. *Measurement and Evaluation in Counseling and Development*, **18**, 3-11.
- Wilson, F. R., Griswold, M. L., & Sunderland, E. (1984). Observer usage in counseling research: An analysis of reporting adequacy. *Measurement and Evaluation in Guidance*, **17**, 40-43.

Manuscript received September 15, 1992; revised February 8, 1993.

This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.