

When the Head Protects the Heart: Empathic Accuracy in Dating Relationships

Jeffrey A. Simpson
Texas A&M University

William Ickes
University of Texas at Arlington

Tami Blackstone
Texas A&M University

This study investigated circumstances in which romantic partners may be motivated to *inaccurately* infer each other's thoughts and feelings. Dating couples rated and discussed pictures of opposite-sex people with whom they might later interact in a dating context. Couples evaluated either highly attractive persons or less attractive persons. As predicted, dating partners who were close, who were insecure about their relationship, and who evaluated highly attractive opposite-sex persons displayed the *least* empathic accuracy when they tried to infer each other's actual thoughts and feelings from the videotape of the rating and discussion task. The effects of these variables were additive, and they were mediated by the degree of perceived threat to the relationship. Theoretical implications of these findings are discussed.

Delusions are as necessary to our happiness as realities.

—Bovee

But love is blind, and lovers cannot see
The pretty follies they themselves commit.

—Shakespeare, *The Merchant of Venice*

Considerable attention has recently focused on empathic accuracy—the extent to which partners in a relationship can accurately infer each other's thoughts and feelings during an interaction episode (Ickes, 1993b). Empathic accuracy has become a focal construct because the degree to which partners can accurately infer what each other is thinking and feeling is presumed to be an indicator of several important facets of a relationship, including its satisfaction, openness, quality of communication, and overall functioning (Knudson, Sommers, & Golding, 1980; Noller & Ruzzene, 1991; Sillars & Scott, 1983).

Recent studies have identified some of the factors that en-

hance empathic accuracy. Relative to couples who avoid conflicts, couples who resolve conflicts in a direct and open fashion tend to develop more accurate understandings of each other's thoughts and feelings as their discussions unfold (Knudson, Sommers, & Golding, 1980), perhaps because direct confrontation of problems permits new and diagnostic information to become available. Individuals also tend to be more accurate at judging the emotionally negative behaviors displayed by their partners than the emotionally positive or neutral ones (Gaelick, Bodenhausen, & Wyer, 1985), an outcome that probably reflects the greater salience and importance of negative information in dyadic interactions. In therapy settings, the empathic accuracy of perceivers tends to improve with increasing exposure to clients, and the rate of improvement can be accelerated by providing perceivers with corrective feedback about clients' actual thoughts and feelings (Marangoni, Garcia, Ickes, & Teng, 1995). Same-sex friends tend to display greater empathic accuracy than same-sex strangers, apparently because they have more detailed, personal knowledge about each other (Stinson & Ickes, 1992). Empathic accuracy in mixed-sex dyads tends to be greater if people are more behaviorally involved in the interaction, if perceivers have higher grade point averages, if targets are more physically attractive, and if perceivers report a larger percentage of partner-relevant thoughts and feelings (Ickes, Stinson, Bissonnette, & Garcia, 1990).

Empathic Accuracy and Relationship Quality

At first glance, the relation between empathic accuracy and relationship quality appears to be strangely inconsistent. Some evidence indicates that empathic understanding and relationship quality are positively associated (e.g., Kahn, 1970; Noller & Ruzzene, 1991; Ruzzene, 1990), whereas other evidence points to a negative link (e.g., Floyd, 1988; Gottman, 1979; Ko-

Jeffrey A. Simpson and Tami Blackstone, Department of Psychology, Texas A&M University; William Ickes, Department of Psychology, University of Texas at Arlington.

We thank our undergraduate research assistants at Texas A&M University—John Engelhardt, Mindy Hall, Dan Hardin, Dorothy Lawrence, Greg Martinez, Yamoria Miller, Ken Santell, Robert Rosenquist, and Christine Wilson—as well as numerous research assistants at the University of Texas at Arlington for their assistance with various phases of this research. We also thank John Finch for his statistical advice and expertise.

Correspondence concerning this article should be addressed to Jeffrey A. Simpson, Department of Psychology, Texas A&M University, College Station, Texas 77843-4235; or to William Ickes, Department of Psychology, University of Texas at Arlington, Arlington, Texas 76019-0528.

walik & Gotlib, 1987; Sillars, Pike, Jones, & Murphy, 1984). Evidence for a positive association includes: (a) Noller and Ruzzene's (1991) finding that happily married couples tend to be more accurate at identifying both the kind of affect experienced by their partners during conflict episodes and their specific goals and intentions; (b) Ruzzene's (1990) finding that people in relatively happy marriages are more accurate in judging the affect and intentions of their spouses; (c) Kahn's (1970) finding that satisfied married couples, relative to dissatisfied couples, have less discrepant interpretations of nonverbal communication during interactions; and (d) clinical observations that high levels of inaccuracy often characterize unhappy, dysfunctional relationships (see Gottman, 1979).

On the other hand, evidence for a negative association between empathic accuracy and relationship quality can also be found. This evidence includes: (a) Sillars and Scott's (1983) observation that, in studies focusing on high-conflict issues, couples who are close and relatively happy display lower levels of empathic accuracy than couples who are less close and relatively unhappy; (b) the finding that greater accuracy concerning partners' assessments of the importance of problems discussed during an interaction is associated with more negative vocal tone and more negative verbal statements (Sillars et al., 1984); and (c) the finding that, compared to happily married couples, unhappy couples tend to agree more closely with neutral observers regarding the extent to which messages conveyed during their interactions are positive versus negative in tone (Floyd, 1988; Gottman, 1979; Kowalik & Gotlib, 1987).

The key to resolving this paradox may lie in an observation advanced by Sillars et al. (1984): The relation between empathic accuracy and relationship quality should be positive when more mundane, less conflictive, or less threatening issues to the relationship are the focus of discussion, but it should be negative when more important, more conflictive, or more threatening issues must be confronted. To state this proposition metaphorically, empathic accuracy is like a sharp-edged tool that can cut both ways. It can have a positive effect on relationship quality when it is used to cut through the relatively trivial misunderstandings and minor disagreements that surface in any close relationship. Conversely, because it can evoke intense negative feelings when it is used to uncover differences that threaten the continued existence of a relationship, empathic accuracy can at times exert a profoundly negative effect on relationship quality—cutting right to the exposed core of the partners' pain and dissatisfaction.

A fascinating implication of this two-edged sword view of empathic accuracy is the possibility that relationship partners might, under some circumstances, be motivated to *inaccurately* infer each others' thoughts and feelings in order to spare themselves—and their relationship—the undesirable consequences that might otherwise result. Unfortunately, however, past research has focused almost exclusively on conditions that enhance empathic accuracy; little, if any, research has systematically explored the conditions associated with decreased accuracy. Accordingly, a major goal of the present study was to examine conditions in which empathic inaccuracy might occur in the context of a close relationship in the service of relationship maintenance.

Empathic Inaccuracy and Relationship Maintenance

Based on a comprehensive review of the interpersonal perception literature, Sillars and Scott (1983) noted that biases in interpersonal perception tend to occur when interactions are highly stressful, when interactants experience strong positive or negative emotions, and when the perceptual task is ambiguous in nature (see also Fletcher & Kininmonth, 1991). In view of these considerations, several theoretical perspectives suggest predictions about the specific conditions in which relationship partners should be most likely to display empathic inaccuracy.

As a rule, people should be motivated to misperceive the thoughts and feelings of their partners if accurate perceptions might threaten the stability of well-established, close relationships (cf. Kelley, 1979). According to interdependence theory (Kelley & Thibaut, 1978; Thibaut & Kelley, 1959), the stability of a relationship depends on the extent to which it provides outcomes better than those that could be found in the best alternative relationship. If dating partners are exposed to highly attractive, available alternative partners (i.e., those capable of raising the comparison level for alternatives), the stability of their relationship should be threatened and they should attempt to defend and maintain it.

Guided by interdependence theory, recent research on relationship maintenance has revealed that people involved in relationships appear to control and diminish their *own* attraction to desirable alternative partners by derogating the alternative partners' personal attributes (Johnson & Rusbult, 1989) and their physical attractiveness (Simpson, Gangestad, & Lerma, 1990). However, no research to date has examined the perceptual processes associated with how people deal with their *partner's* attraction to alternative partners during actual dyadic interactions. According to a relationship maintenance perspective, individuals should make less accurate inferences about what their partner is thinking and feeling when he or she evaluates highly attractive alternatives because accurate inferences might destabilize the relationship, especially if the partner is harboring strong, positive thoughts and feelings about the desirability of the alternatives. Accordingly, we predicted that when dating partners evaluate and discuss alternative partners, couples who discuss highly attractive alternatives should display *less* empathic accuracy concerning each other's actual thoughts and feelings than couples who discuss less attractive alternatives.

When interactions focus on alternative partners, certain features of the relationship in which partners are involved should produce heightened perceptions of threat. The development of closeness in a relationship requires the investment of considerable time, effort, and resources (Kelley et al., 1983), and the dissolution of a close relationship often generates substantial emotional distress (Berscheid, 1983; Simpson, 1987). Because of this, partners who have a relatively close relationship (i.e., a relationship in which they have strong, frequent, and diverse impact on each other over time: see Kelley et al., 1983) should be more threatened by the potentially destabilizing effects of alternative partners. As a consequence, partners involved in close relationships should be more strongly motivated to shield themselves from the threats posed by alternative partners than should partners involved in relationships that are less close. Ac-

cordingly, we predicted that couples involved in close relationships, relative to those who are less close, would display *less* empathic accuracy when evaluating and discussing alternative partners.

Perceived threat also should be greater among partners who are insecure about the long-term stability of their relationship (cf. Berscheid & Fei, 1977). Whereas relationship closeness reflects the extent to which an individual has strong, frequent, and diverse emotional and behavioral ties to a relationship partner, insecurity reflects the degree to which an individual is confident that his or her partner is strongly bound to the relationship. When couples who are insecure must discuss and evaluate alternative partners, the vulnerability and tenuous nature of their relationship should be accentuated. To diminish this strong and unavoidable threat, insecure partners ought to *inaccurately* infer the thoughts and feelings that their partner may be having about alternative partners, because accurate assessments might be too threatening and perhaps damaging to the relationship. Thus, we predicted that insecure couples would display *less* empathic accuracy when discussing alternative partners than secure couples would.¹

All three conditions—exposure to attractive alternatives, greater closeness, and greater insecurity—should be associated with greater perceived threat and, therefore, with less empathic accuracy. Perceived threat and empathic inaccuracy should be most pronounced when all three conditions are simultaneously present in a relationship. Hence, we predicted that couples who are close, insecure, and exposed to attractive alternatives would display the *least* empathic accuracy, particularly when compared to couples possessing the opposite constellation of features.

It is important to emphasize that the biases underlying empathic inaccuracy are likely to differ from person to person. In response to threat, some people may display less empathic accuracy because they underestimate the degree to which their partner perceives opposite-sex alternatives to be appealing. Other people may misdirect their attention toward features of alternative partners that are less desirable than those to which their partner is really attending. Still others might avoid making any inferences that their partner is having thoughts or feelings about alternative partners. Although each of these responses reflects a different source of empathic inaccuracy, they all should serve the same function: to reduce the negative impact of strong, external threats to the current relationship. Thus, we offer no predictions about what specific kinds of misperceptions should play a larger role in reducing empathic accuracy, and we do not address this issue in the present investigation.

Another major unresolved issue concerns how empathic accuracy relates to long-term relationship stability. Marital partners who have incongruent perceptions of each other's attitudes, role expectations, and self-perceptions tend to have poorly adjusted marriages (Dymond, 1954; Ferguson & Allen, 1978; Murstein & Beck, 1972), and poor adjustment is a predictor of divorce (Gottman, 1979). Most previous research, however, has focused on the extent to which partners have concordant impressions of each other in nonthreatening situations. When a direct and unavoidable threat to the relationship must be handled, empathic accuracy should not necessarily predict greater

relationship stability considering the potentially damaging effects that accurate inferences might have.

However, relationship stability might be associated with the level of *disparity* in empathic accuracy that exists between partners. High levels of disparity should be indicative of a power imbalance in the relationship (Ickes, 1993a; 1993b), and this imbalance should have deleterious effects on the long-term stability of the relationship from the vantage point of both the more accurate partner and the less accurate partner. If one partner has substantially more insight into the other partner's thoughts and feelings, the less accurate partner may feel and indeed may be more vulnerable to cognitive, emotional, and behavioral manipulation. Conversely, the more empathically accurate partner may feel misunderstood—a situation that could produce resentment and feelings of inequity in the relationship. In addition, pronounced disparity in empathic accuracy should make it more difficult for partners to feel that they share (or eventually can share) important experiences with each other, and this also might destabilize relationships (Ickes, 1993a; 1993b). Based on these considerations, we hypothesized that dating partners who exhibit more similar (i.e., less discrepant) levels of empathic accuracy should have relationships that are more likely to endure over time.

In sum, this study tested two major hypotheses: (a) Dating couples who are faced with a clear, unavoidable threat to their relationship (i.e., having to evaluate attractive, opposite-sex persons as potential dates), who have a close relationship, or who are insecure about the future of their relationship, should display less empathic accuracy during dyadic interaction. Empathic inaccuracy should be greatest in relationships that possess all three attributes, and least in relationships that possess none of them; (b) Dating partners who exhibit smaller disparities in empathic accuracy during the laboratory interaction should be more likely to remain together over time.

Method

Overview

Dating couples evaluated the physical and sexual attractiveness of a group of men and women who ostensibly were participants in a local "dating pool." In Phase 1, the male and female members of 82 dating couples completed questionnaires that assessed the closeness of their relationship, their level of insecurity about its permanence, and other relationship-based measures. In Phase 2, they jointly viewed a series of slides during which the male partner rated 6 prospective female dates according to their physical attractiveness and sexual appeal, and the female partner rated 6 prospective male dates on the same dimensions. After stating their numerical ratings aloud, the partners were asked to discuss each stimulus person for 30 s. The couples viewed either highly

¹ Recent research indicates that two major dimensions underlie romantic relationships: the extent to which partners are close and interdependent and the extent to which they are emotionally insecure (Simpson, Hebl, & Nations, 1994). The Relationship Closeness Inventory (RCI; Berscheid, Snyder, & Omoto, 1989) is a central marker of the first dimension, and the Insecurity Scale is a strong marker of the second one. Thus, measures tapping each of these dimensions were used in the present investigation.

attractive stimulus persons or less attractive ones. While they rated and discussed the slides, the couples were unobtrusively videotaped.

During Phase 3, the male and female partners independently viewed the videotape in separate rooms, indicating at what points during the interaction they experienced a thought or feeling, what it was, and whether it was positive, neutral, or negative. Following this, each partner was asked to view the interaction again, stopping the tape at each point when their *partner* reported having had a specific thought or feeling. At each stop point, the participant was asked to infer what her or his partner had been thinking or feeling. Empathic accuracy scores were later computed from these data. Four months later, during Phase 4, the dating partners were contacted by telephone to determine whether they were still dating.

Participants

Eighty-two heterosexual dating couples (82 women and 82 men) participated in a study on attraction and relationships. At least one member of each couple was enrolled in introductory psychology at Texas A&M University and received partial course credit for her or his participation.

Procedure

Phase 1. During Phase 1, the male and female members of each couple responded to a large battery of self-report questionnaires in private. The participants were assured that their partner would not be allowed to see their answers, and they were asked to answer all questions as accurately and honestly as possible. Embedded in the survey were two measures gathered specifically for this study: (a) the Relationship Closeness Inventory (the RCI; Berscheid, Snyder, & Omoto, 1989), and (b) the Insecurity Scale (Fei & Berscheid, 1977). Available data support the reliability and validity of both of these measures (see Berscheid et al., 1989, and Fei & Berscheid, 1977).² In the present sample, the internal consistencies for the full RCI (summed across all three components) and the Insecurity Scale were .62 and .80, respectively. Participants also indicated how long they had dated each other (in months) and whether they were dating each other exclusively.

Phase 2. Once both members of each couple had completed the questionnaire, they were reunited and led to an experimental room containing a table, a slide projector, two chairs, and a viewing screen. They were asked to sit in the two chairs, which faced the viewing screen. The chairs were situated next to one another, approximately 18 inches (.5 m) apart. The experimenter then read the following instructions:

The next phase of this study involves rating photographs of individuals who have agreed to take part in a dating study. We already have several personality measures from our volunteers, but we need ratings of their physical attractiveness and sexual appeal. We are asking dating couples to make these ratings because you, in effect, are experts on dating. Due to the delicate nature of arranging dates, we want to be as confident of our matches as possible. Later this semester, we may ask you to conduct interviews with our dating study volunteers. In the event you are asked, we would arrange for you to privately interview one of your two most highly rated individuals.

The experimenter emphasized that it was critical to the success of the study for both partners to provide honest, accurate ratings of each stimulus person, based on each partner's personal perceptions. The couples then viewed a series of 12 slides depicting 6 male and 6 female undergraduates. Couples randomly assigned to the high-threat condition saw 12 individuals who were highly attractive. In reality, these individuals were college-age professional models. Couples randomly assigned to the low-threat condition saw 12 individuals who were some-

what less than average in attractiveness. The pictures of these individuals were taken from a college yearbook. Within each condition, the order of viewing (male partners rating slides of women first vs. female partners rating slides of men first) was counterbalanced. Because preliminary analyses revealed that the order in which ratings were made did not affect the results, this variable will not be discussed further.

The male and female dating partners rated all opposite-sex persons on the dimensions of physical attractiveness and sexual appeal by stating both ratings aloud, after which they discussed what they liked and disliked about each person with their dating partner. Specifically, when the slide depicting each stimulus person appeared on the screen, the experimenter asked the opposite-sex dating partner, "How physically attractive do you find this person to be as a potential dating partner?" and "How sexually appealing do you find this person to be as a potential dating partner?" Both questions were answered aloud using 10-point scales, anchored by 1 (*not at all*) and 10 (*extremely*). Following the study, we aggregated the 12 ratings (the 6 physical attractiveness ratings and the 6 sexual appeal ratings made by each dating partner) to create a single global index of the stimulus persons' rated physical attractiveness and sexual appeal. This global attractiveness index was extremely reliable, with Cronbach alphas of .96 and .95, respectively, for the male and female dating partners.

When the appropriate partner had made the ratings, the experimenter recorded them and asked the couple for a brief discussion of the ratings. The couple then proceeded to the next slide. Each slide remained on the screen for 30 s. Once the initial task instructions had been given, the entire procedure took approximately 6 min. During this time, each couples' interaction was covertly recorded by a concealed videocamera mounted in the corner of the room.

When the rating task was finished, the experimenter had the participants complete a three-item manipulation check measure that assessed how threatened they were by the stimulus persons their partners had evaluated. Using a 9-point Likert-type scale anchored by 1 (*not at all*) and 9

² Available evidence supports the reliability and validity of both the RCI and the Insecurity Scale. In terms of reliability, each of the subscales of the RCI are internally consistent, and the coefficient alpha for the full RCI index is .62 (Berscheid et al., 1989). The RCI also has high test-retest reliability over 3-5 week intervals ($r = .82$). In terms of convergent and predictive validity, the RCI distinguishes between close and less close relationships (Berscheid et al., 1989), it correlates with several face-valid indicators of closeness (Berscheid et al., 1989), and it predicts both long-term relationship stability (Berscheid et al., 1989; Simpson, 1987) and the extent of postdissolution emotional distress (Simpson, 1987). As for discriminant validity, the RCI is only moderately correlated with measures of liking, loving, subjective closeness, and the frequency with which various emotions are experienced in the relationship (Berscheid et al., 1989).

The Insecurity Scale has high internal consistency ($\alpha = .91$). With regard to convergent validity, Fei and Berscheid (1977) have marshaled evidence that people who score higher on the Insecurity Scale report less current life satisfaction ($r = -.33$), score higher on the Janis-Field Feelings of Inadequacy Scale (Hovland & Janis, 1959; $r = .28$), and believe that their current romantic partners like them less ($r = -.35$), love them less ($r = -.27$), and are less committed to the relationship ($r = -.25$). Highly insecure individuals also display more distrust in their relationships, and people tend to score higher on the Insecurity Scale when reporting on serious relationships versus more casual ones. In terms of discriminant validity, scores on the Insecurity Scale do *not* correlate with measures of intelligence, scores on the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1964), dependency on the partner, or respondents' age, birth order, gender, or parental income.

(*extremely*), participants indicated the degree to which they felt threatened, jealous, and upset while their partner did the rating task. Responses to these three items were aggregated to form a composite manipulation check measure of perceived threat (Cronbach alpha = .90).

All couples were then told that they had been videotaped. The experimenter explained that they were not informed about the videotaping prior to the study because such knowledge might have led them to interact in an unnatural or atypical way. The dating partners were then asked to give their informed consent for their videotapes to be released for research purposes. It was made clear that if either of them did not want the tape used for any reason, they could erase it themselves immediately. No couple chose to do so. Participants then were informed that the individuals they rated were not volunteers for a dating study, and that neither partner would be asked to meet and privately interview one of their two most highly rated individuals. We provided participants with this debriefing information so they would focus their attention on the Phase 3 thought and feeling tasks. We reasoned that participants would provide more honest and accurate data in Phase 3 if they had been fully debriefed and were not thinking about an impending interaction with one of the stimulus persons.

Phase 3. In Phase 3, the participants were led to separate rooms that each contained a VCR and a color monitor. Here, they independently viewed the videotape of their interaction. Following procedures used in previous empathic accuracy studies (e.g., Stinson & Ickes, 1992), participants were instructed to provide an accurate, honest, and complete account of all the thoughts and feelings they had during the rating task. In addition, participants were guaranteed that their partners would not see their actual thought and feeling entries. These procedures were designed to increase the likelihood that the thoughts and feelings reported by both partners were accurate and uncensored.

The participants then viewed their interaction twice. During the first viewing, each partner was asked to watch the tape and to stop it when he or she remembered having had a specific thought or feeling at that moment in the interaction. Following a procedure described by Ickes et al. (1990), participants were instructed to: (a) write down each thought or feeling on a standardized form that was provided; (b) indicate the number of minutes and seconds into the interaction when each thought or feeling occurred (using a timer that was displayed on the VCR); (c) specify whether each entry was a thought or a feeling; and (d) state whether each entry was positive, neutral, or negative in its overall emotional tone. The participants were again assured that their partners would not be shown any of their responses. The experimenter stayed in the room long enough to establish that each participant followed the procedure correctly, at which point she left. The men recorded a mean of 12.7 thought and feeling entries (range: 5–30), and the women recorded a mean of 11.8 entries (range: 2–28).

After both partners had viewed the videotape once, they were given a second set of thought/feeling forms similar to the first ones. Although this set contained each of the times at which *their partner* had reported having a specific thought or feeling, it did not contain the contents of those thoughts or feelings. The participants' task during this second viewing of the tape was to: (a) record what they thought their partner had been thinking or feeling at each of the times specified; (b) indicate whether each inferred entry was a thought or a feeling; and (c) rate or judge its overall emotional tone.

Following this assessment, both partners were thanked for their participation and more fully debriefed. As in the earlier debriefing, they were reminded that the individuals they rated were not part of a dating pool, and that the study concerned the extent to which dating partners could accurately infer each other's thoughts and feelings when exposed to attractive versus less attractive opposite-sex persons. The reason for the unobtrusive videotaping was carefully explained, and participants were told that their videotapes would be seen only by trained research

assistants. No couple was allowed to leave until the experimenter was convinced that both partners felt good about participating in the study and fully understood why various procedures were followed.

Phase 4. Approximately 4 months later, an attempt was made to recontact all participants by telephone. Both members of 79 dyads (96% of the original sample) were successfully contacted. Participants responded to a telephone survey that inquired, among other things, about whether they were still dating the same partner. In particular, they were asked: "When you participated in the study last fall, you were dating a person named [name of dating partner]. Are you still dating this person?" (answered either "Yes" or "No"). Both partners had to agree about the current status of their relationship (still dating vs. no longer dating) for their data to be used on this follow-up measure. Only one couple did not agree about the current status of their relationship.

Construction of the Empathic Accuracy and Covariate Measures

Two sets of measures were constructed from the participants' ratings of the videotapes. The first set included the original and baseline measures of empathic accuracy from which the final, revised measure of empathic accuracy was computed. The second set included measures of two theoretically relevant covariates: the level of stress associated with each thought or feeling reported by each participant during the slide rating task (rated by four independent judges), and the relative ease versus difficulty of inferring each thought or feeling from each participant's verbal and nonverbal behavior on the videotape (also rated by four independent judges). As discussed in greater detail below, the two covariate measures were collected to potentially rule out two alternative explanations: (a) that the lower empathic accuracy of partners who were close, insecure, and exposed to attractive alternatives was attributable to higher levels of stress or tension in their thoughts and feelings compared to other types of partners; or (b) that their lower empathic accuracy was attributable to the fact that their thoughts and feelings were more difficult to infer.

Empathic accuracy measures. Empathic accuracy was operationally defined as the degree to which a dyad member's written description of the inferred content of his or her partner's thought/feeling entry matched the actual content of the sentence the partner had written when reporting the thought or feeling. All of the actual and inferred thought/feeling entries were typed into word processor files that provided the input to one of the programs in the COLLECT YOUR THOUGHTS software package (Ickes et al., 1990). When this program was run, each dyad member's actual thought/feeling entry appeared in sequence on the top portion of the CRT screen, with the corresponding inferred entry appearing directly below it. Separate copies of the software (and the accompanying thought and feeling data) were given to six independent judges who had been trained individually in its use. Each judge compared the written content of each actual entry with that of the corresponding inferred entry and rated the degree of similarity (i.e., content accuracy) using a 3-point scale ranging from *essentially different content* (0), through *somewhat similar, but not the same, content* (1), to *essentially the same content* (2).³

With the six independent judges treated as items, the internal consistency of this original measure of empathic accuracy was .93. A second

³ An examination of the thoughts and feelings of couples in the high-threat condition revealed few instances in which one partner inferred greater attraction to the stimulus person than was apparent in the other partner's actual thoughts or feelings about the stimulus person. Thus, misperceptions of this nature probably did not play a major role in reducing empathic accuracy.

software program was then used to compute the mean of the six judges' similarity ratings for each actual-compared-with-inferred thought/feeling entry. These means were summed across all of the actual/inferred entry pairs within a given participant's protocol. The summed values were then divided by the maximum number of accuracy points that could be obtained for a given number of actual/inferred entry pairs in order to derive an overall accuracy score that controlled for individual differences in the number of entries/inferences reported. This last step resulted in a global accuracy score that was conveniently rescaled so that it had a possible range of .00 (total inaccuracy) to 1.00 (perfect accuracy).

Following the logic and procedures described by Ickes et al. (1990), a measure of baseline accuracy was similarly derived. Another program in the COLLECT YOUR THOUGHTS software package was used to randomly pair actual and inferred thought/feeling entries. A different set of six independent judges made similarity ratings of these randomly paired actual/inferred entries using the same software program described above. With the six independent judges treated as items, the internal consistency of this baseline measure of empathic accuracy was .83. A revised measure of empathic accuracy was then computed by subtracting the baseline accuracy score for each participant (i.e., perceiver) from the original global accuracy score for that participant. This adjusted measure of empathic accuracy was used as the dependent variable in the analyses reported below. It reflects the empathic accuracy of each perceiver that remains after removing a baseline component which estimates the degree of accuracy that might be expected by chance.

Stress measure. Four independent judges then reviewed all of the actual thought/feeling entries and rated each one according to whether it contained stress or tension. Ratings were made on a 2-point scale such that 0 = *no stress* and 1 = *stress*. The interrater reliability computed across all of the entries was .85. The judges' ratings for the stress content of each thought/feeling entry were then aggregated and divided by the total number of thought/feeling entries reported by each participant to create a global stress index for each partner.

Inferential difficulty measure. Finally, four independent judges who were not involved in the stress rating task watched each couple's videotaped interaction and stopped the tape at each point when either the male or the female partner reported having a thought or feeling. They then read what the thought or feeling was and rated it according to how difficult it would be for a perceiver (i.e., the partner) to accurately infer its content. Ratings were made on a 2-point scale such that 0 = *relatively easy to infer this thought or feeling* and 1 = *relatively difficult to infer this thought or feeling*. The judges' ratings for each thought/feeling entry were then aggregated for each partner. Across all entries, the interrater reliability was .97. We then divided the judges' ratings by the total number of thought/feeling entries made by each participant to create a global inferential difficulty index for each partner.

Results

Because the dyad members' responses were reliably correlated (i.e., nonindependent) for nearly all of the measures used in this study, we averaged the responses of the male and female partners on each measure to create aggregated, dyad-level variables (see Kenny, 1988; Kenny & LaVoie, 1985). Except for some repeated-measures analyses in which the male and female partners' scores are treated as levels of a within-dyad variable, all of the results reported below are dyad-level findings in which the couple serves as the unit of analysis. It is important to note that the two nonmanipulated independent variables—relationship closeness and insecurity—were almost perfectly orthogo-

nal in this sample, regardless of whether they were treated as continuous measures, $r = -.08$, *ns*, or as dichotomous ones, $r = -.05$, *ns*.

Manipulation Checks

We hypothesized that the level of perceived threat during the rating task would be greater for couples who (a) evaluated highly attractive alternative partners, (b) were close, and (c) were insecure about their relationship. To test this prediction, we conducted a 2 (relationship closeness: high vs. low) \times 2 (insecurity: high vs. low) \times 2 (threat level of slides: high vs. low) \times 2 (gender: male vs. female) repeated-measures analysis in which closeness, insecurity, and threat were treated as between-dyad factors and gender was treated as a within-dyad factor. The dependent measure was each partner's score on the three-item measure of perceived threat.

As expected, three main effects emerged. Couples who evaluated more attractive stimulus persons reported more perceived threat during their discussions than did couples who evaluated less attractive stimulus persons, $M_s = 20.6$ and 15.3 , respectively, $F(1, 74) = 5.58$, $p < .03$. Couples who scored higher on the RCI reported feeling more threat than did couples who scored lower, $M_s = 21.3$ and 16.4 , respectively, $F(1, 74) = 6.27$, $p < .02$. And couples who were more insecure reported more threat than couples who were less insecure, $M_s = 21.6$ and 16.0 , respectively, $F(1, 74) = 7.32$, $p < .01$. These effects were additive, such that perceived threat was greatest in couples who were close, insecure, and who evaluated highly attractive alternatives. There were no interaction effects.

As an additional check on the success of our manipulated threat variable, we examined participants' aggregated ratings of the physical attractiveness and sexual appeal of the 6 stimulus persons whom they evaluated during the slide rating task. Specifically, we conducted a 2 (relationship closeness: high vs. low) \times 2 (insecurity: high vs. low) \times 2 (threat level of slides: high vs. low) \times 2 (gender: male vs. female) repeated-measures analysis in which closeness, insecurity, and threat were treated as between-dyad factors and gender was treated as a within-dyad factor. The dependent measure was each partner's rating of the six stimulus persons whom he or she evaluated (i.e., each partner's score on the global physical and sexual attractiveness index). Only one main effect emerged. Stimulus persons in the high-threat condition were rated as significantly more physically attractive and sexually appealing than stimulus persons in the low threat condition, $F(1, 74) = 214.00$, $p < .0001$. Breaking the data down by gender indicated that the male participants who viewed the more attractive set of women rated them higher ($M = 6.5$) than did men who viewed the less attractive women ($M = 3.0$), $t(80) = 15.10$, $p < .0001$. Similarly, female participants who evaluated the more attractive men rated them higher ($M = 6.3$) than did women who evaluated the less attractive men ($M = 3.1$), $t(80) = 12.65$, $p < .0001$. No effects emerged for either closeness or insecurity, both $F_s < 2.20$, *ns*.

Because lower empathic accuracy could result from systematic distortions in the mean ratings of stimulus persons provided by certain couples (particularly those who were close, insecure, and who evaluated attractive stimulus persons), we also con-

ducted a series of planned contrasts on the physical/sexual attractiveness index. Specifically, we tested whether the mean attractiveness ratings provided by couples who were close, insecure, and who saw attractive stimulus persons ($M = 6.3$) differed from the means of couples in the other high-threat conditions (i.e., high threat/less closeness/less insecurity: $M = 6.3$; high threat/more closeness/less insecurity: $M = 6.0$; high threat/less closeness/more insecurity: $M = 6.6$). No significant differences were found, all t s < 1, *ns*. Thus, compared with other couples who also were assigned to the high-threat condition, couples who were closer and more insecure neither underestimated nor overestimated the attractiveness of the stimulus persons whom they rated. Any empathic inaccuracy displayed by this group, therefore, probably does not stem from systematic distortions in the ratings of the stimulus persons.

Results for the Empathic Accuracy Measure

To test our primary hypotheses, we first conducted a 2 (relationship closeness: high vs. low) \times 2 (insecurity: high vs. low) \times 2 (threat level of slides: high vs. low) \times 2 (gender: male vs. female) repeated-measures analysis in which closeness, insecurity, and threat were treated as between-dyad factors and gender was treated as a within-dyad factor. The dependent measures were the male and female partners' scores on the empathic accuracy index. Three main effects and a single two-way interaction emerged (see Table 1).⁴

As the means in Table 1 indicate, couples in the high-threat condition displayed less empathic accuracy than did couples in the low-threat condition, $F(1, 71) = 4.25, p < .05$. Couples who scored higher on the RCI showed less empathic accuracy than couples who scored lower on it, $F(1, 71) = 3.98, p < .05$. And couples who exhibited greater insecurity were less accurate than couples who were more secure, $F(1, 71) = 5.22, p < .03$. The only significant interaction found was a Gender \times Threat interaction, which indicated that women were more empathically accurate than men in the low-threat condition, whereas men were more accurate than women in the high-threat condition, $F(1, 71) = 9.86, p < .002$. Tests for gender differences within each condition revealed that the women displayed greater empathic accuracy ($M = 16.7$) than the men ($M = 10.7$) in the low-threat condition, $t(54) = 2.09, p < .05$. However, in the high-threat condition, the men were marginally more accurate ($M = 11.6$) than the women ($M = 7.8$), $t(100) = 1.81, p < .08$.

Planned Contrasts

We next conducted a series of planned contrasts (Kirk, 1982). Given that (a) none of the planned contrasts involving empathic accuracy were qualified by interactions with gender, and (b) the male and female partners' empathic accuracy scores were nonindependent ($r = .20, p < .15$) using the criterion recommended by Kenny and LaVoie (1985), we averaged the partners' scores on the empathic accuracy index and conducted three planned contrasts on this aggregated (i.e., dyad-level) measure of empathic accuracy.

The first contrast pitted the empathic accuracy scores of couples who were high in closeness, high in insecurity, and exposed

Table 1
Means and Standard Deviations for the Dyad-Level
Empathic Accuracy Index

Relationship status	Situational threat			
	High		Low	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
High closeness/high insecurity	5.1 ^a	7.9	11.9 ^b	6.9
High closeness/low insecurity	11.5 ^c	9.6	9.9 ^d	5.2
Low closeness/high insecurity	8.7 ^e	7.8	13.7 ^f	7.6
Low closeness/low insecurity	13.8 ^g	7.9	17.6 ^h	9.4

Note. Number of dyads per cell: ^a = 13; ^b = 5; ^c = 13; ^d = 6; ^e = 13; ^f = 9; ^g = 12; ^h = 8.

to highly attractive opposite-sex persons ($M = 5.1$) against the average empathic accuracy scores of all other couples in the study ($M = 12.3$). This contrast revealed that couples who were highly threatened, close, and insecure were reliably less accurate during the rating task than couples who did not have this specific configuration of attributes, $t(77) = 2.87, p < .01$.

The second contrast compared couples who were close, insecure, and situationally threatened with those who possessed the opposite set of attributes (i.e., couples who were less close, less insecure, and less threatened). This contrast revealed that highly threatened, close, and insecure couples displayed markedly less empathic accuracy ($M = 5.1$) than did couples with the opposite set of attributes, $M = 17.6, t(19) = 3.41, p < .01$.

In a previous study, Ickes et al. (1990) examined the empathic accuracy of opposite-sex strangers in an unstructured, weak situational context. Given that the same measure of empathic accuracy was used in both the Ickes et al. study and the present study, the empathic accuracy of opposite-sex strangers provides a reference point against which the extent of empathic inaccuracy in the present investigation can be gauged. Accordingly, the third contrast compared the empathic accuracy scores of the high-threat/high-closeness/high-insecurity couples in the present study against the

⁴ We also conducted a regression analysis in which the dyad-level measure of empathic accuracy was treated as the dependent measure, and situational threat (dichotomously coded), closeness (assessed by continuous scores on the RCI), insecurity (measured by continuous scores on the Insecurity Scale), and all two- and three-way interactions served as predictors. Results nearly identical to those reported above were found when regression techniques were used. Although closeness and insecurity were analyzed as continuous measures in most of the analyses described below (e.g., the Probit analyses and the mediation analyses), closeness and insecurity were dichotomized, and repeated measures analyses were conducted initially for two reasons. First, we wanted to compare the level of empathic accuracy in various conditions of this study with the level of empathic accuracy observed in previous studies involving strangers, all of which have reported and analyzed cell means for empathic accuracy (see Ickes et al., 1990). Comparisons across different studies are necessary to gauge the relative level of empathic accuracy that is occurring in a given social context. Second, the use of a repeated measures analysis strategy was needed in order to conduct the hypothesized planned contrasts.

average empathic accuracy scores of all of the mixed-sex dyads in the Ickes et al. (1990) study. This contrast revealed that the high-threat/high-closeness/high-insecurity couples displayed reliably less empathic accuracy than did mixed-sex strangers in an unstructured interaction setting, $M_s = 5.1$ and 15.9 , respectively, $t(49) = 3.35, p < .01$.

Alternative Explanations

We predicted that couples who are highly close, insecure about their relationship, and confronted with strong situational threat should be most likely to display empathic inaccuracy in order to insulate their relationships from destabilizing inferences. Yet empathic inaccuracy might exist for reasons other than defensive, motivationally based ones. For example, perceivers may have difficulty accurately inferring the thoughts and feelings of others if the verbal and nonverbal behaviors that others display are difficult to read or interpret. If couples who are close, insecure, and highly threatened provide each other with behavioral cues that are difficult to interpret, the empathic inaccuracy we observed may not stem from motivated inaccuracy per se. Similarly, if such couples differ from other couples in the level of stress or tension that is evident in their thought and feeling entries, their empathic accuracy could be attenuated because higher stress might either distract their attention from the verbal and nonverbal cues emitted by their partners or interfere with complex inferential activity (cf. Schroder, Driver, & Streufert, 1967; Suedfeld & Tetlock, 1977).

To examine these possibilities, we conducted two analyses of covariance in which relationship closeness (high vs. low), insecurity (high vs. low), and threat level of slides (high vs. low) were between-dyad factors, gender (male vs. female) was a within-dyad factor, and each partner's score on the empathic accuracy index served as the dependent variable. In the first analysis, each partner's score on the difficulty-of-inference index was treated as the covariate; in the second analysis, each partner's score on the stress index served as the covariate.

When scores on the difficulty-of-inference index were statistically controlled, two of the three main effects along with the single two-way interaction reported above were still significant or marginally significant. Specifically, couples who evaluated more threatening (i.e., attractive) stimulus persons still displayed less empathic accuracy compared to couples who evaluated less threatening persons, $F(1, 69) = 5.72, p < .02$. Couples who were more insecure continued to be less empathically accurate than couples who were more secure, although this effect was marginal, $F(1, 69) = 3.72, p < .06$. The third main effect for closeness was slightly more attenuated, $F(1, 69) = 2.07, p < .16$. The Gender \times Threat interaction—in which women were more empathically accurate than men in the low-threat condition, but men were more accurate than women in the high-threat condition—remained reliable, $F(1, 69) = 9.42, p < .003$.⁵

When scores on the stress index were controlled for, two of the three main effects and the lone two-way interaction remained significant or marginally significant. Couples who were more insecure continued to exhibit less empathic accuracy, $F(1, 67) = 4.88, p < .04$. Couples who were closer displayed marginally less accuracy, $F(1, 67) = 3.06, p < .09$. And couples who eval-

uated more attractive stimulus persons were less accurate, although this effect was marginal, $F(1, 67) = 2.74, p = .10$. The Gender \times Threat interaction also remained reliable, $F(1, 67) = 6.81, p < .02$.⁶

The Gender \times Threat Interaction: A Closer Look

To clarify the nature of the Gender \times Threat interaction for empathic accuracy, we examined its relationship with the two covariates: participants' difficulty of inferring the thoughts and feelings reported by their dating partners and the level of stress evident in their own thoughts and feelings during the rating task. Accordingly, we conducted two additional 2 (relationship closeness: high vs. low) \times 2 (insecurity: high vs. low) \times 2 (threat level of slides: high vs. low) \times 2 (gender: male vs. female) repeated-measures analyses in which closeness, insecurity, and threat were between-dyad factors and gender was the within-dyad factor. In the first analysis, the partners' scores on the difficulty-of-inference index were the dependent measures; in the second analysis, the partners' scores on the stress index were the dependent measures.

When the difficulty-of-inference index was the dependent measure, a marginally significant main effect emerged for situational threat, indicating that the partners' thoughts and feelings were judged by raters as somewhat more difficult to infer in the high-threat condition than in the low-threat condition, $F(1, 69) = 2.87, p < .10$. The Gender \times Threat interaction, however, was not significant, $F(1, 69) < 1$, suggesting that difficulty-of-inferences was not a plausible mediator of this interaction effect.

When the stress index was the dependent measure, couples who were assigned to the high-threat condition, $F(1, 67) = 9.94, p < .005$, and who were closer, $F(1, 67) = 6.19, p < .02$, reported thoughts and feelings that reflected more stress or tension, as judged by raters. In addition, the Gender \times Threat interaction was significant, $F(1, 67) = 7.58, p < .01$. The pattern of means indicated that women experienced somewhat less stress ($M = 7.4$) than did men ($M = 11.3$) in the low-threat condition, whereas women experienced slightly more stress ($M = 14.2$) than men ($M = 12.4$) in the high-threat condition. These findings suggest that the stress level of the perceiver was a possible mediator of the Gender \times Threat interaction effect.

⁵ Correlational analyses revealed that female partners were less empathically accurate when the thoughts and feelings of their male partners were rated as more difficult to infer, $r = -.32, p < .01$. There was no significant association, however, between the empathic accuracy displayed by the male partners and the extent to which the thoughts and feelings of their female partners were difficult to infer, $r = -.10, ns$. These correlations were not significantly different in magnitude, $z = 1.42, ns$.

⁶ Similar to the inferential difficulty analyses, female partners were less empathically accurate when the thoughts and feelings of their male partners were rated as containing greater stress, $r = -.24, p < .05$. However, no association emerged between empathic accuracy displayed by male partners and the extent to which the thoughts and feelings of their female partners were rated as containing stress, $r = .07, ns$. The difference between these correlations proved to be marginally significant, $z = 1.93, p < .10$.

Viewed as a pattern, these results indicate that the decline in empathic accuracy experienced by women in the high-threat condition did *not* occur because the thoughts and feelings of their male partners were more difficult to infer. Although women in the high-threat condition did evidence greater stress, it is important to note that the Gender \times Threat interaction remained reliable even when participants' level of stress was statistically controlled. Our subsequent analyses, therefore, centered on the most plausible and theoretically interesting mediator: the degree of perceived threat reported by dating partners.

Mediation Analyses

We hypothesized that the effects of the three independent variables on empathic accuracy should be mediated by the level of perceived threat that participants felt during the rating and discussion task. To test for this mediation, we performed a series of path analyses using the LISREL 7 program (Joreskog & Sorbom, 1989). The exogenous variables in these analyses were the dyad-level measures of closeness, insecurity, and experimentally manipulated threat.⁷ The endogenous variables in the first set of analyses were the dyad-level measures of perceived threat and empathic accuracy. When perceived threat was *not* allowed to serve as the mediator between the three independent variables and empathic accuracy, all of the paths from the independent variables to empathic accuracy were significant or marginally significant, but the model provided a rather poor fit to the data, $\chi^2(4) = 30.98$, $p < .0001$, goodness-of-fit index = .88, Adjusted goodness-of-fit index = .57. As indicated in Figure 1, when perceived threat was introduced as the mediator, all paths were significant and the fit improved considerably, $\chi^2(3) = 6.36$, *ns*, goodness-of-fit index = .97, Adjusted goodness-of-fit index = .85. We then calculated the chi square value for the null model in which all paths were set to 0. The normed fit index (the NFI: Bentler & Bonett, 1980) indicated that the mediation model fit the data fairly well, NFI = .86.

Additional analyses were conducted to determine whether either the inferential difficulty index or the stress index mediated relations between the three independent variables and empathic accuracy. Path analyses using LISREL 7 revealed that the inferential difficulty index was not a good mediator, $\chi^2(3) = 11.13$, $p < .02$, goodness-of-fit index = .95, adjusted goodness-of-fit index = .75. The stress index, however, was a fairly good mediator ($\chi^2[3] = 5.33$, *ns*, goodness-of-fit index = .97, adjusted goodness-of-fit index = .87), providing a much better fit compared to when it was not allowed to serve as a mediator, $\chi^2(4) = 25.17$, $p < .0001$, goodness-of-fit index = .90, adjusted goodness-of-fit index = .63. Next, we calculated the null model and found that the Bentler and Bonett index confirmed this conclusion, NFI = .88.

Given these findings, we then tested whether perceived threat mediated relations between the three independent variables and empathic accuracy, even when participants' level of stress was statistically controlled. To accomplish this, we partialled scores on the stress index from the relations between empathic accuracy, perceived threat, and the three independent variables. When the effect of stress was statistically controlled, the fit of the model in which perceived threat was treated as a mediator

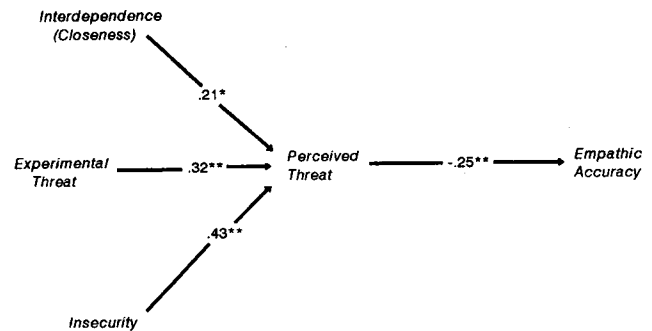


Figure 1. Standardized betas are presented. * $p < .05$. ** $p < .01$.

actually improved slightly, $\chi^2(3) = 4.33$, *ns*, goodness-of-fit index = .98, adjusted goodness-of-fit index = .89. Although definitive conclusions cannot be drawn from these analyses, they are consistent with the hypothesis that perceived threat mediates the links between the three independent variables and empathic accuracy, and that it does so independently of stress. These results suggest that perceived threat and stress may both be important mediating variables that reduce empathic accuracy via different routes.

Relationship Stability

Four months after the laboratory task, both partners of each dyad were contacted by telephone and asked whether they were still dating one another. Of the 78 couples who were reached and who agreed about the status of their relationship, 60 couples were still dating and 18 had disbanded. We first conducted a Probit analysis in which relationship stability (coded 1 = *still dating* and 0 = *no longer dating*) was the dependent measure, and situational threat, closeness (the continuous dyad-level measure), and insecurity (the continuous dyad-level measure) served as predictors. The results revealed that couples who were closer, $F(1, 73) = 5.54$, $p < .01$, and who were less insecure, $F(1, 73) = 3.69$, $p < .10$, were more likely to be dating at follow-up, although the effect for insecurity was marginally significant. No other significant main effects or interactions emerged.

We next tested the hypothesis that couples who displayed more similar levels of empathic accuracy should be less likely to break up. Preliminary analyses indicated that among couples who were still dating, the correlation between the empathic accuracy scores of the male and female partners was positive and reliable, $r = .29$, $p < .05$. For couples who had broken up, there was a nonsignificant, negative correlation between the partners' empathic accuracy scores, $r = -.10$, *ns*. To test whether greater similarity predicted long-term relationship stability, we conducted another Probit analysis following recommendations for

⁷ Because the level of threat in the laboratory task was dichotomous (i.e., high or low), we calculated point-biserial correlations between threat and the other continuous measures prior to conducting the path analyses. This allowed us to estimate what the relations between threat and the other measures would have been if threat had been a continuous measure.

testing discrepancy scores proposed by Cohen and Cohen (1983). Specifically, we entered the three independent variables (threat, closeness, and insecurity) as a block at Step 1, the female partner's empathic accuracy score at Step 2, and the male partner's empathic accuracy score at Step 3. If the variable entered in the last step (male empathic accuracy) significantly predicts relationship stability when the preceding predictor variable (female empathic accuracy) is partialled out, the discrepancy between these two measures is significant. The results revealed that male scores on the empathic accuracy index significantly predicted relationship stability when entered in the final step, $F(1, 73) = 4.12, p < .05$. Thus, partners who were more *similar* to each other in empathic accuracy during the laboratory interaction were more likely to be dating 4 months later.

To determine whether the absolute *level* of empathic accuracy forecasted relationship stability, we also conducted a Probit analysis in which relationship stability was the dependent variable and the dyad-level measure of empathic accuracy was the predictor. No effect emerged, $F < 1, ns$. In fact, the point-biserial correlation between relationship stability and absolute level of empathic accuracy was $r = .00$. Hence, couples who were more empathically accurate during the laboratory task were neither more nor less likely to disband compared to couples who exhibited less empathic accuracy.

None of the 13 couples in the high situational-threat/high-closeness/high-insecurity condition had broken up at follow-up, whereas the breakup rate for couples in the other conditions averaged 28%. To determine whether the breakup rate was significantly less among couples in the high-threat/high-closeness/high-insecurity condition compared to couples in all other conditions, we conducted a chi-square test in which the occurrence of dissolution for couples constituting each "group" was the dependent measure (see Hays, 1981). The results indicated that the breakup rate was significantly lower for couples in the high-threat/high-closeness/high-insecurity condition compared to all other couples, $\chi^2(1) = 4.68, p < .05$. Thus, in line with a motivated inaccuracy interpretation, couples who were closer, more insecure, and who evaluated more attractive stimulus persons were relatively *more* likely than other couples to still be dating 4 months later.

Discussion

In the present study, we investigated conditions in which dating partners might be motivated to *inaccurately* infer each other's thoughts and feelings. Dating partners who were close, insecure about their relationship, and required to evaluate highly attractive opposite-sex persons in each other's presence displayed the *least* empathic accuracy when they attempted to infer each other's thoughts and feelings. These effects were mediated by the degree to which partners felt threatened during the rating task. The mediational effect of perceived threat on empathic accuracy was not attributable to the amount of stress evident in dyad members' thoughts and feelings or to the ease with which their thoughts and feelings could be inferred from their behavior. Finally, partners who displayed similar (as opposed to divergent) levels of empathic accuracy in the rating

task were more likely to be dating 4 months later, whereas the absolute level of couples' empathic accuracy was unrelated to long-term relationship stability.

The Case for Motivated Inaccuracy

Researchers have only begun to explore the conditions that promote empathic inaccuracy in relationships. The present findings suggest one set of conditions that may generate motivated inaccuracy in romantic relationships. Although it is premature to claim that the case for motivated inaccuracy is definitive, six lines of argument are consistent with this interpretation. First, in addition to the nonmanipulated variables of relationship closeness and relationship insecurity, the experimentally manipulated variable of situational threat produced decrements in empathic accuracy. According to a motivated inaccuracy perspective, strong situational threat should act as the primary motivational impetus that triggers perceptions of threat which, in turn, should impair empathic accuracy. The results of our path analysis confirm each of the conjectured links in our model.

Second, our experimental procedures were designed to minimize the likelihood that empathic inaccuracy could have resulted from participants' censoring or failing to report their actual thoughts and feelings about the stimulus persons. Couples were explicitly instructed to be as accurate as possible, both in publicly making their attractiveness ratings of the stimulus persons and in privately recording their thoughts and feelings. For couples in the high-situational-threat condition, the mean attractiveness ratings made by couples who should have been most daunted by the rating task (i.e., couples who were close and insecure) did not differ from the mean attractiveness ratings made by other couples. Similarly, for couples in the low-situational-threat condition, the mean attractiveness ratings did not differ between any of the cells. Decrements in empathic accuracy, therefore, do not seem to be a result of systematic distortions in the ratings made by different types of couples. A more plausible explanation seems to be motivated inaccuracy, which locates the source of empathic inaccuracy in the perceiver's biased inferences rather than in the partner's biased reporting.

Third, among couples in the high-situational-threat condition, there were relatively few instances in which one partner inferred greater attraction to the depicted stimulus person than was apparent in the other partner's actual thoughts or feelings. Misperceptions in the direction of overestimated attraction, therefore, were not a significant source of empathic inaccuracy. Clearly, misperceptions should not heighten perceived threat if motivated inaccuracy is operating.

Fourth, if motivated inaccuracy is occurring, perceived threat should mediate the relations between empathic accuracy and the independent variables used in this study. Our path analyses confirmed this mediation. They also revealed that the mediational effect of perceived threat was independent of both the level of stress apparent in dating partners' thoughts or feelings and the ease with which their thoughts or feelings could be accurately inferred. Hence, dating partners who were closer, more insecure, and faced with strong situational threat were not more

inaccurate simply because they experienced greater stress or because their partners' thoughts and feelings were more difficult to infer. In accord with a motivated inaccuracy perspective, their greater inaccuracy derived in large part from heightened perceptions of threat.

Fifth, the follow-up data indicated that the subset of couples who displayed the *least* empathic accuracy in the laboratory task were *more* likely than other couples to still be dating 4 months later. This finding suggests that when accurate inferences about a partner's current thoughts or feelings might destabilize or undermine an existing relationship, greater empathic inaccuracy may actually enhance relationship stability. The potential benefits of empathic inaccuracy, however, may be limited to circumstances in which threat to a relationship is strong and cannot be easily ignored, averted, or discounted.

Sixth, the empathic accuracy exhibited by couples who were closer, more insecure, and exposed to highly attractive others is substantially lower than that found in *any* of the previous empathic accuracy studies, despite the fact that most of these studies examined the empathic accuracy of total strangers (see Ickes, 1993b). Why should well-acquainted dating partners display significantly less empathic accuracy than opposite-sex strangers unless they were motivated to inaccurately perceive each other's thoughts and feelings? Considering all of the available evidence, the most plausible and parsimonious explanation appears to be that dating partners were motivated to disregard, misinterpret, or distort the content of each other's actual thoughts and feelings as a function of the perceived threat to their relationship.

Two Mediators of Empathic Inaccuracy in Women?

An unanticipated Gender \times Threat interaction effect emerged for the empathic accuracy measure. Women displayed greater empathic accuracy than men when couples evaluated less attractive alternative partners, but men displayed slightly greater accuracy than women when highly attractive alternatives were rated. This interaction was driven primarily by the greater empathic accuracy of women in the low-threat condition relative to women in the high-threat condition. Subsequent analyses within the high-threat condition revealed that women did not report more perceived threat than men. Moreover, women in the high-threat condition did not have dating partners whose thoughts and feelings were more difficult to infer compared to women in the low-threat condition, yet their thoughts and feelings did convey more stress.

These findings suggest that decrements in empathic accuracy for men were predominately in response to perceived threat, whereas decrements for women were *independently* affected by both perceived threat and stress. That is, two independent factors seem to have produced the lower level of empathic accuracy displayed by women in the high-threat condition. Besides the impairing effects of perceived threat, heightened stress can reduce empathic accuracy by drawing perceivers' attention away from the verbal and nonverbal cues emitted by partners or by impeding complex inferential thought processes after cues have been encoded (Schroder et al., 1967; Suedfeld & Tetlock, 1977).

Why were higher levels of stress evident in the thoughts and feelings of women in the high-threat condition? One possibility is that women might have found the rating task to be more awkward and uncomfortable. Relative to men, women usually take a more active role in maintaining romantic relationships (cf. Surra & Longstreth, 1990). Because of this, women may be less inclined to openly express their views about desirable alternative partners in the presence of their current partner, and they may feel more uncomfortable doing so. These conjectures await future research.

Relationship Stability and Similarity in Empathic Accuracy

The relationship stability results revealed that dating partners who displayed more *similar* (i.e., positively correlated) levels of empathic accuracy were more likely to still be dating at 4-month follow-up. Why should similarity in empathic accuracy enhance relationship stability? Several possibilities exist. First, to the extent that the perception of equity strengthens most relationships (see Backman & Secord, 1966; Walster, Berscheid, & Walster, 1973), pronounced inequities in empathic accuracy could destabilize relationships by creating corresponding inequities in feeling understood by one's partner, feeling vulnerable to emotional exploitation by one's partner, or feeling that different levels of commitment or investment exist in the relationship (Ickes, 1993a; 1993b). Second, disparities in empathic accuracy might reflect flawed intersubjectivity in relationships, a situation in which partners cannot bring their individual understandings into synchrony with each other (Ickes, 1993b; Ickes & Simpson, in press). Both of these processes should contribute to greater instability in the relationships of partners with dissimilar levels of empathic accuracy.

Another major question for future research involves whether, or the extent to which, psychological mechanisms that produce motivated inaccuracy operate at the level of conscious awareness. It is conceivable that the selective information processing which takes place in conditions of high threat may occur outside of conscious awareness. Because perception is governed by cognitive processes that operate at different levels of awareness yet are under the same central regulative control, Erydeli (1974) has suggested that many psychological processes associated with perceptual defense and vigilance *should* occur at subliminal levels of awareness. Using computer simulations, Hinton (1992) has recently modeled how perceptual selectivity could operate subconsciously at the level of neural networks within the brain. Evidence that motivated inaccuracy stems from subconscious processes would lend support to the proposition that motivated inaccuracy operates as a psychological mechanism serving a primary defensive function: to sidetrack or minimize conscious perceptions of threat to an existing relationship.

Conclusions and Caveats

The present study has a number of features that limit the scope of its conclusions. First, although the extent of threat was manipulated to be either high or low, evaluating opposite-sex persons according to their physical and sexual attractiveness in

front of one's dating partner should have been an awkward and relatively stressful experience, even for couples who evaluated less attractive persons. The effects reported above, therefore, may be confined to stressful contexts. Second, because two independent variables could not be manipulated (relationship closeness and relationship insecurity), causal conclusions apply to only the manipulated independent variable—situational threat. Due to ethical and practical considerations, we did not attempt to manipulate relationship attributes. Third, it is not possible to determine whether empathic inaccuracy occurred because individuals were motivated to protect themselves (e.g., their self-esteem or self-worth) or the stability of their relationship. Indeed, individuals may have experienced threat from both sources, which are often naturally confounded (see Aron, Aron, Tudor, & Nelson, 1991). These caveats notwithstanding, the present investigation opens new and exciting avenues for future research on the origins, the nature, and the multifaceted functions that motivated inaccuracy may serve in romantic relationships.

References

- Aron, A., Aron, E. N., Tudor, M., & Nelson, G. (1991). Close relationships as including others in the self. *Journal of Personality and Social Psychology*, *60*, 241–253.
- Backman, C. W., & Secord, P. F. (1966). The compromise process and the affect structure of groups. In C. W. Backman & P. F. Secord (Eds.), *Problems in social psychology* (pp. 190–192). New York: McGraw-Hill.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606.
- Berscheid, E. (1983). Emotion. In H. H. Kelley, E. Berscheid, A. Christensen, J. H. Harvey, T. L. Huston, G. Levinger, E. McClintock, L. A. Peplau, & D. R. Peterson, *Close relationships* (pp. 110–168). San Francisco: Freeman.
- Berscheid, E., & Fei, J. (1977). Romantic love and sexual jealousy. In G. Clanton & L. G. Smith (Eds.), *Jealousy* (pp. 101–109). Englewood Cliffs, NJ: Prentice-Hall.
- Berscheid, E., Snyder, M., & Omoto, A. M. (1989). The Relationship Closeness Inventory: Assessing the closeness of interpersonal relationships. *Journal of Personality and Social Psychology*, *57*, 792–807.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Crowne, D., & Marlowe, D. (1964). *The approval motive*. New York: Wiley.
- Dymond, R. (1954). Interpersonal perception and marital happiness. *Canadian Journal of Psychology*, *8*, 164–171.
- Erydeli, M. H. (1974). A new look at the new look: Perceptual defense and vigilance. *Psychological Review*, *81*, 1–25.
- Fei, J., & Berscheid, E. (1977). *Perceived dependency, insecurity, and love in heterosexual relationships: The eternal triangle*. Unpublished manuscript, University of Minnesota.
- Ferguson, L. R., & Allen, D. R. (1978). Congruence of parental perception, marital satisfaction, and child adjustment. *Journal of Consulting and Clinical Psychology*, *46*, 345–346.
- Fletcher, G. J. O., & Kininmonth, L. (1991). Interaction in close relationships and social cognition. In G. J. O. Fletcher & F. D. Fincham (Eds.), *Cognition in close relationships* (pp. 235–255). Hillsdale, NJ: Erlbaum.
- Floyd, F. J. (1988). Couples' cognitive/affective reactions to communication behaviors. *Journal of Marriage and the Family*, *50*, 523–532.
- Gaelick, L., Bodenhausen, G. V., & Wyer, R. S. (1985). Emotional communication in close relationships. *Journal of Personality and Social Psychology*, *49*, 1246–1265.
- Gottman, J. M. (1979). *Marital interaction: Experimental investigations*. New York: Academic Press.
- Hays, W. L. (1981). *Statistics* (3rd ed.). New York: Holt, Rinehart & Winston.
- Hinton, G. E. (1992). How neural networks learn from experience. *Scientific American*, *267*, 144–151.
- Hovland, C., & Janis, I. (1959). *Personality and persuasibility*. New Haven, CT: Yale University Press.
- Ickes, W. (1993a, June). Closeness and mutual understanding. In H. Reis & C. Rusbult (Chairs), *What is closeness?* Symposium conducted at the 4th conference of the International Network on Personal Relationships, Milwaukee, Wisconsin.
- Ickes, W. (1993b). Empathic accuracy. *Journal of Personality*, *61*, 587–610.
- Ickes, W., & Simpson, J. A. (in press). Managing empathic inaccuracy in close relationships. In W. Ickes (Ed.), *Empathic accuracy*. New York: Guilford.
- 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.
- Johnson, D. J., & Rusbult, C. E. (1989). Resisting temptation: Devaluation of alternative partners as a means of maintaining commitment in close relationships. *Journal of Personality and Social Psychology*, *57*, 967–980.
- Joreskog, K. G., & Sorbom, D. (1989). *Lisrel 7: A guide to the program and applications* (2nd ed.). Chicago: SPSS, Inc.
- Kahn, M. (1970). Nonverbal communication and marital satisfaction. *Family Processes*, *9*, 449–456.
- Kelley, H. H. (1979). *Personal relationships: Their structure and process*. Hillsdale, NJ: Erlbaum.
- Kelley, H. H., Berscheid, E., Christensen, A., Harvey, J., Huston, T. L., Levinger, G., McClintock, E., Peplau, L. A., & Peterson, D. R. (1983). Analyzing close relationships. In H. H. Kelley, E. Berscheid, A. Christensen, J. Harvey, T. L. Huston, G. Levinger, E. McClintock, L. A. Peplau, & D. R. Peterson, *Close relationships* (pp. 20–67). San Francisco: Freeman.
- Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence*. New York: Wiley.
- Kenny, D. A. (1988). The analysis of data from two-person relationships. In S. Duck (Ed.), *Handbook of personal relationships* (pp. 57–78). London: Wiley.
- Kenny, D. A., & LaVoie, L. (1985). Separating individual and group effects. *Journal of Personality and Social Psychology*, *48*, 339–348.
- Kirk, R. E. (1982). *Experimental design* (2nd ed.). Monterey, CA: Brooks/Cole.
- Knudson, R. M., Sommers, A. A., & Golding, S. L. (1980). Interpersonal perception and mode of resolution in marital conflict. *Journal of Personality and Social Psychology*, *38*, 751–763.
- Kowalik, D. L., & Gotlib, I. H. (1987). Depression and marital interaction: Concordance between intent and perception and communication. *Journal of Abnormal Psychology*, *96*, 127–134.
- Marangoni, C., Garcia, S., Ickes, W., & Teng, G. (1995). Empathic accuracy in a clinically relevant setting. *Journal of Personality and Social Psychology*, *68*, 854–869.
- Murstein, B. I., & Beck, G. D. (1972). Person perception, marriage adjustment, and social desirability. *Journal of Consulting and Clinical Psychology*, *39*, 396–403.
- Noller, P., & Ruzzene, M. (1991). Communication in marriage: The influence of affect and cognition. In G. J. O. Fletcher & F. D. Fincham

- (Eds.), *Cognition in close relationships* (pp. 203–233). Hillsdale, NJ: Erlbaum.
- Ruzzene, M. (1990). *Attributional processes in marriage: Judgments of cognition, affect, and causality in conflict interactions*. Unpublished manuscript, University of Queensland, Australia.
- Schroder, H. M., Driver, M. J., & Streufert, S. (1967). *Human information processing*. New York: Holt, Rinehart & Winston.
- Sillars, A. L., Pike, G. R., Jones, T. S., & Murphy, M. A. (1984). Communication and understanding in marriage. *Human Communication Research, 10*, 317–350.
- Sillars, A. L., & Scott, M. D. (1983). Interpersonal perception between intimates: An integrative review. *Human Communication Research, 10*, 153–176.
- Simpson, J. A. (1987). The dissolution of romantic relationships: Factors involved in relationship stability and emotional distress. *Journal of Personality and Social Psychology, 53*, 683–692.
- Simpson, J. A., Gangestad, S. W., & Lerma, M. (1990). Perception of physical attractiveness: Mechanisms involved in the maintenance of romantic relationships. *Journal of Personality and Social Psychology, 59*, 1192–1201.
- Simpson, J. A., Hebl, M., & Nations, C. (1994). *Trust and interdependence in romantic relationships*. Unpublished manuscript, Texas A&M University.
- 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.
- Suedfeld, P., & Tetlock, P. (1977). Integrative complexity of communications in international crises. *Journal of Conflict Resolutions, 21*, 169–184.
- Surra, C. A., & Longstreth, M. (1990). Similarity of outcomes, interdependence, and conflict in dating relationships. *Journal of Personality and Social Psychology, 59*, 501–516.
- Thibaut, J. W., & Kelley, H. H. (1959). *The social psychology of groups*. New York: Wiley.
- Walster, E., Berscheid, E., & Walster, G. W. (1973). New directions in equity research. *Journal of Personality and Social Psychology, 25*, 151–176.

Received September 20, 1993

Revision received October 3, 1994

Accepted April 6, 1995 ■