Journal of Personality and Social Psychology
1997, Vol. 73, No. 4, 747–757

0022-3514/97/$3.00

Confidence and Accuracy in Person Perception: Do We Know What We Think We Know About Our Relationship Partners?

William B. Swann, Jr., and Michael J. Gill
University of Texas at Austin

A cross-sectional study of dating partners and a longitudinal study of college roommates revealed that the confidence and accuracy of their impressions were often dissociated. For example, relationship length and degree of involvement tended to increase the confidence of people's impressions, but neither variable consistently increased the accuracy of their impressions of their partners' sexual histories, activity preferences, and so on. A third study showed that relationship length and involvement increased the richness of impressions, and richness fostered confidence. The authors conclude that although confidence–accuracy dissociations are surely problematic in some instances, their apparent pervasiveness raises the possibility that confidence may sometimes contribute to relationship quality even when it is unrelated to accuracy.

"Can I entrust my neighbor with the key to my house?" "Can I be sure that the job candidate will be as industrious as her reference letters suggest?" "Is my daughter's new boyfriend as earnest as he seems?" The confidence with which people answer such questions is critically important, for confidence serves as a psychological gatekeeper of sorts, systematically determining whether people translate their beliefs into action (e.g., Berger & Mitchell, 1989; Fazio & Zanna, 1978; Pieters & Verplanken, 1995; Swann & Ely, 1984). And perhaps this is as it should be. After all, people presumably become confident of their beliefs because they possess sound evidence for these beliefs, and sound evidence, in turn, presumably fosters accuracy. Highly confident beliefs should therefore be highly accurate.

But are they? The research literature provides surprisingly little support for the notion that confidence is closely tied to accuracy. Instead, confidence and accuracy often seem to be dissociated. Some of the most direct evidence of confidence–accuracy dissociations comes from studies in which clinicians grew progressively more confident in their impressions of clients while the accuracy of their impressions remained low (Oskamp, 1965; Ryback, 1967). In addition, studies of eyewitness testimony have revealed virtually no relation between the confidence and accuracy of people's beliefs (e.g., Wells & Murray, 1984). Even when people make judgments about "objective facts" (e.g., the capital of Sweden), there seems to be little relation between the confidence and accuracy of their beliefs (see Jacoby, Bjork, & Kelley, 1994, for a review).

William B. Swann, Jr., and Michael J. Gill, Department of Psychology, University of Texas at Austin.

We thank Melissa Plesh for her assistance in the empirical phase of Study 3, and Daniel Gilbert, Brett Pelham, and Bobbie Spellman for commenting on an earlier version of this article. This research was supported by National Science Foundation Grant SBR-9319570.

Correspondence concerning this article should be addressed to William B. Swann, Jr., Department of Psychology, University of Texas, Austin, Texas 78712. Electronic mail may be sent via the Internet to swann@psy.utexas.edu.

To be sure, confidence and accuracy are sometimes linked. Dunning, Griffin, Milojkovic, and Ross (1990), for example, discovered that people's confidence in predicting specific behaviors of others *was* related to the accuracy of those predictions. Although the confidence–accuracy relations reported by Dunning et al. (1990) were quite modest, they raise the possibility that confidence and accuracy may be associated under some circumstances.

Which circumstances? We propose that to answer this question, one must first specify why confidence and accuracy are sometimes dissociated.

Confidence–Accuracy Dissociations: Why?

We propose that as people become increasingly intimate with others, the richness of their impressions increases. There are two reasons for such increases. First, they acquire more information. Second, they become motivated to integrate what they know into coherent impressions. Therefore, for both informational and motivational reasons, over time people develop a wealth of tightly integrated information about their relationship partners. Such rich impressions are relatively accessible (Sherman & Klein, 1994; Smith, Adams, & Schorr, 1978) and should, in turn, promote confidence (Kelley & Lindsay, 1993; Koriat, 1993; Nelson & Narens, 1990).

Although rich representations of others may foster confidence, they will not necessarily contribute to accuracy. The problem is that nondiagnostic information may contribute to richness just as much as diagnostic information (e.g., Jacoby et al., 1994; Kelley & Lindsay, 1993; Koriat, 1993). For example, learning that one's relationship partner has certain qualities (e.g., "gives to charity," "exudes innocence," "is a regular churchgoer") could make one more confident of an objectively unrelated quality (e.g., "has not had many sexual partners") if one (either implicitly or explicitly) erroneously assumes that all such qualities are expressions of "morality." In such instances, nondiagnostic information will foster richness and confidence but not accuracy.

To explore the relation between confidence and accuracy among everyday person perceivers, we examined the perceptions of couples who had been dating for varying lengths of time and who expressed varying levels of involvement with their partners. We reasoned that compared to people in new or uninvolved relationships, people in lengthy or highly involved relationships would have more opportunities to gather information about one another, more motivation to acquire information (Berscheid, Graziano, Monson, & Dermer, 1976), and more motivation to integrate that information into coherent representations (Murray & Holmes, 1993). Thus, a lengthy, highly involved relationship should increase the richness of one's representation of one's partner. Although these increments in *representational richness* should increase the confidence of people's impressions, they will increase the accuracy of their impressions only insofar as the information is diagnostic. Unfortunately, the information relationship partners acquire may not be all that diagnostic. For example, research suggests that accuracy increases early in our relationships but improves little, if at all, after that (e.g., Albright, 1990; Kenny, 1994). Hence, as the amount of time and involvement in a relationship increase, people's impressions may become increasingly confident but no more accurate. They may thus develop an illusory feeling of knowing their partner.

Study 1: Confidence and Accuracy Among Dating Couples

Method

Participants

We recruited 80 heterosexual dating couples by placing an advertisement in a University of Texas newspaper and by inviting the participation of students in Introductory Psychology. Our sample included 63 couples who responded to the newspaper ad and 17 couples from the psychology classes. Couples received either $15 or credit in their course for participating.

Participants had been dating for 3–312 weeks, with an average relationship length of 78 weeks. They ranged in age from 17 to 41, with a mean age of 21. When asked, 15% described themselves as Asian, 4% as Black, 20% as Hispanic, 55% as White, and 6% as other.

Because Bernieri, Zuckerman, Koestner, and Rosenthal (1994) suggested that cohabitation moderates accuracy, and because there were relatively few cohabiting couples, we eliminated 22 couples who reported living together.[1] In addition, we excluded 11 couples from analyses involving sexual history, because they were not sexually active and thus had no variance on these items. Finally, 1 couple was excluded because of missing data. This left 57 couples for most of the analyses and 46 couples for analyses involving sexual history.

Procedure

On arrival at the laboratory, members of each couple were randomly assigned to the roles of target (who made self-ratings) and perceiver (who rated his or her partner). Same-sex experimenters escorted perceivers and targets to separate rooms. Perceivers learned that we were interested in "how accurately dating partners know one another." Targets, in contrast, were introduced to a study of "the relation between personality and relationship satisfaction/commitment." We misled targets about our objectives because we suspected that they would be more candid if they did not know the true purpose of our study (e.g., they would refrain from trying to remember what they told their partner about themselves). Also, we assured all participants that their partners would not see their responses.

Targets completed a sexual history questionnaire (SHQ), based on Metzler, Noell, and Biglan's (1992) measure of high-risk sexual behavior. The SHQ requires respondents to answer 10 open-ended questions concerning their number of past sexual partners, frequency of condom usage, frequency of discussing sexual history before beginning a sexual relationship, average number of dates before engaging in intercourse, sexually transmitted disease history, and so on. Targets also completed the Self-Liking/Competence Scale (SLC; Tafarodi & Swann, 1995) and the Self-Attributes Questionnaire (SAQ; Pelham & Swann, 1989). The SLC is a 20-item measure tapping both the "liking" and "competence" components of global self-esteem, and the SAQ is a 15-item measure tapping the respondent's self-perceived intelligence, social competence, artistic/musical ability, athleticism, and attractiveness; the certainty with which the respondent holds these five self-views; and the importance of these self-views to the respondent. Finally, targets completed the Activity Preferences Questionnaire (APQ; Surra & Longstreth, 1990), a 37-item measure that requires respondents to rate the extent to which they enjoy various activities (e.g., cleaning their rooms, going to a bar, playing board games, etc.).[2]

Perceivers completed the same questionnaires as the targets but were instructed to "answer each questionnaire the way you think your partner would answer."[3] For each item (or each cluster of related items on the APQ and SLC) perceivers also rated how confident they were (from 0–100%) that they accurately predicted their partner's response to that item. In addition, for each of our measures (e.g., the APQ) perceivers estimated the number of items on which their predicted response exactly matched their partner's response.

Targets and perceivers both completed an 8-item relationship involvement questionnaire (Rusbult, 1980) that required them to indicate how long they would like their relationship to last, how much they loved their partner, how their relationship compared to their ideal, and so on. Finally, targets and perceivers reported demographic information and indicated how long they had been dating each other.

Computation of Confidence Indices

We created confidence indices for each of the scales by computing the average of perceivers' confidence ratings on each questionnaire (αs = .87, .84, and .91 for the SHQ, APQ, and SAQ, respectively; the SLC contained a single confidence rating only). As can be seen in Table 1, perceivers varied widely in their levels of confidence; just as some perceivers had serious doubts about the accuracy of their predictions, others were certain that their predictions were accurate.

Computation of Accuracy Indices

The literature on accuracy in person perception suggests two distinct ways to assess accuracy (e.g., Bernieri et al., 1994). *Within-dyad accu-*

[1] None of the conclusions were substantially altered when cohabiting couples were added to the analyses.

[2] These measures include numerous items that tap constructs featured in the Big 5 taxonomy. For example, our activity preference and self-concept measures tap Extraversion, our self-esteem measure contains items that tap the inverse of the Neuroticism factor, and so on.

[3] We asked perceivers to respond as they thought targets would to ensure that perceivers and targets were making inferences about the same things. In principle, this strategy should maximize the accuracy of perceivers relative to a procedure in which perceivers estimate what targets are really like. Other research suggests, however, that when perceivers are asked to estimate the target's "true qualities" rather than his or her questionnaire responses, our effects involving confidence and accuracy would remain unchanged (Gill, Swann, & Silvera, 1997).

racy (sometimes called *profile accuracy*) is computed within dyads across questionnaire items. This commonly used index of person-perception accuracy measures the extent to which a perceiver can predict the pattern of a target's responses across the items (e.g., Andersen, 1984; Bernieri, 1991; Bernieri et al., 1994; Snodgrass & Rosenthal, 1985). Another common index of accuracy is *between-dyad accuracy*, which is computed across dyads on individual traits (e.g., Funder, 1995; Kenny, 1994). This index measures the extent to which perceivers can predict the scores of targets on a single-trait item or on a questionnaire that measures a single construct.

Because each of these approaches to measuring person perception accuracy has advantages and disadvantages (for a discussion, see Bernieri et al., 1994), we used both of them. We should note, however, that we believe that the within-dyad accuracy index is particularly appropriate in the present context. One reason is that the within-dyad index creates an accuracy score for each perceiver that can be correlated with perceiver confidence, whereas the between-dyad index provides only a single accuracy coefficient that refers to the accuracy of "the group." In addition, the between-dyad accuracy approach may overestimate accuracy in studies like ours. Because the content of some of our questionnaires varied widely across items, perceivers may predict targets' average scores on the questionnaires accurately while doing a poor job of discriminating the patterns of targets' responses within questionnaires. For example, if a target reports having had many sexual partners and having never contracted an STD, a perceiver could appear accurate at predicting the target's average score even if he or she predicted that the target had had few sexual partners and many STDs. Although this problem is minimized to the extent that a questionnaire has a quite high (rather than merely adequate) level of internal consistency, the internal consistency of our measures is rarely quite high.

Computing within-dyad accuracy. Our measures of within-dyad accuracy were intraclass correlations (Shrout & Fleiss, 1979) computed within each couple on each of our four questionnaires. So computed, the intraclass correlation assesses the extent to which a target's and a perceiver's ratings show a similar trend across items. For example, on the SAQ a high intraclass correlation would result when both the target and perceiver agree that the target is high in intellectual ability and social competence but low in artistic/musical ability, athleticism, and physical attractiveness. Our intraclass correlations were based on a two-way mixed effect analysis of variance in which rater (i.e., target vs. perceiver) was treated as a fixed effect, and questionnaire item (e.g., SHQ Item 1 vs. SHQ Item 2 vs. SHQ Item 3, etc.) was treated as a random effect. The numerator of our intraclass correlation formula was the difference between the mean squares associated with the items effect and the error mean squares, and the denominator was equal to the mean squares for items plus two times the error mean squares. This is formula ICC(3, 1) in Shrout and Fleiss (1979).

There was considerable variability among couples when we examined their levels of within-dyad accuracy. In fact, the data in Table 2 reveal that the range of accuracy correlations was equal to the entire theoretical range of the intraclass correlation. Also, when we repeated these analyses using Pearson correlation coefficients, the mean levels of accuracy were virtually identical (as were the relations between accuracy and our independent variables).[4]

Table 1
Descriptive Statistics for Confidence Indices (Study 1)

Measure	*M*	*SD*	Range
SHQ	77.7	17.1	40–100
SLC	81.8	13.7	50–100
SAQ	78.9	11.3	39–98
APQ	80.5	11.0	40–97

Note. All numbers are percentages. SHQ = sexual history questionnaire; SLC = Self-Liking/Competence Scale; SAQ = Self-Attributes Questionnaire; APQ = Activity Preferences Questionnaire.

Table 2
Descriptive Statistics for Within-Dyad Accuracy Correlations (Study 1)

Measure	*M*	*SD*	Range
SHQ	.51	.35	.00–1.0
SLC	.59	.37	.00–1.0
SAQ	.52	.26	.00–.85
APQ	.57	.14	.15–.84

Note. SHQ = sexual history questionnaire; SLC = Self-Liking/Competence Scale; SAQ = Self-Attributes Questionnaire; APQ = Activity Preferences Questionnaire.

Computing between-dyad accuracy. We computed between-dyad accuracy across couples using Pearson correlations. We began by calculating an average score for each target on each of our four questionnaires (αs = .72, .93, .77, and .79 for the SHQ, SLC, SAQ, and APQ, respectively). High scores on the SHQ represent relatively risky sexual behavior, high scores on the SLC represent relatively high self-esteem, high scores on the SAQ represent relatively positive specific self-views, and high scores on the APQ represent a relative tendency to enjoy all activities. Next we calculated an average score for each perceiver's predictions on each questionnaire (αs = .73, .95, .79, and .79 for the SHQ, SLC, SAQ, and APQ, respectively). Finally, we calculated between-dyad accuracy for each questionnaire by correlating targets' and perceivers' average scores. The resulting correlations were $r(44) = .75$, $p < .001$ for the SHQ; $r(55) = .66$, $p < .001$ for the SLC; $r(55) = .59$, $p < .001$ for the SAQ; and $r(55) = .20$, $p = .13$ for the APQ.

Results and Discussion

Relation of Confidence and Accuracy

Our primary interest was in the relation between the confidence and accuracy of people's impressions. We examined this relation using both within- and between-dyad accuracy approaches. In both cases, each of our four questionnaires was analyzed separately.

[4] Computing a correlation within dyads creates an accuracy measure that contains both "stereotype accuracy" and "true accuracy" (see Cronbach, 1955). To estimate the levels of stereotype accuracy in our sample, we used a "pseudodyad" technique pioneered by Corsini (1956). We randomly assigned target persons to opposite-sex perceivers and computed intraclass correlations within these pseudodyads on each of our four measures. These correlations told us the extent to which a target's responses were related to the predictions of a randomly selected opposite-sex perceiver, and they provide one estimate of stereotype accuracy. To improve the reliability of our estimates of stereotype accuracy, we randomly assigned targets to perceivers two times and took the average of the resulting intraclass correlations as our estimate of stereotype accuracy. Estimates of stereotype accuracy were .24, .49, .28, and .28 for the SHQ, SLC, SAQ, and APQ, respectively. Within-subject t tests revealed that stereotype accuracy was less than total accuracy on all four of our measures (all ps $< .001$). Thus, the accuracy correlations in our sample seem to reflect true accuracy as well as stereotype accuracy.

Confidence and within-dyad accuracy. We began by computing the correlation between perceiver confidence and our within-dyad accuracy correlations. The correlations were quite small for three of our four measures, including the SHQ, $r(44) = .17$, *ns*; the SAQ, $r(55) = -.09$, *ns*; and the APQ, $r(55) = .16$, *ns*, suggesting that people's confidence had little to do with the accuracy of their knowledge about their partner. Although confidence *was* related to accuracy on the SLC, $r(55) = .53$, $p < .01$, on three of four questionnaires confidence and accuracy were dissociated.

Confidence and between-dyad accuracy. We examined whether perceiver confidence was related to the between-dyad accuracy index. We tested this prediction using moderated multiple regressions (Bernieri et al., 1994; Paunonen & Jackson, 1985; Saunders, 1956), a technique for assessing the moderating impact of a continuous variable (e.g., confidence) on the correlation between two other continuous variables (e.g., target self-ratings and perceiver predictions). In this case, we attempted to predict targets' self-ratings from (a) perceivers' predictions, (b) perceivers' confidence, and (c) the interaction of perceivers' predictions and confidence. Confidence can be said to moderate between-dyad accuracy if the interaction term of this regression equation is statistically significant.

Tests of the interaction beta weights revealed that confidence did not moderate between-dyad accuracy on the SLC or APQ (*t*s < 0.6, *ns*),[5] and confidence was negatively associated with between-dyad accuracy on the SAQ, $t(53) = -2.3$, $p < .03$ ($\beta = -.25$). Nevertheless, confidence was positively associated with between-dyad accuracy on the SHQ, $t(42) = 2.3$, $p < .03$ ($\beta = .23$).

Although the moderating influence of confidence on SHQ between-dyad accuracy was significant, the effect size was small ($r^2 = 5\%$) and only slightly stronger than the nonsignificant relation between confidence and within-dyad accuracy ($r^2 = 3\%$). The same was true of the effect size for the negative relation between confidence and between-dyad accuracy on the SAQ ($r^2 = 5\%$). For all practical matters, then, it appears that confidence and between-dyad accuracy are largely unrelated and, when they are correlated, the relation may sometimes be negative.

Relation of Relationship Length and Involvement to Confidence and Accuracy

We expected that relationship length and involvement would produce increases in confidence but not accuracy. We examined confidence first.

Confidence. We first regressed the confidence indices for the SHQ, SLC, SAQ, and APQ on relationship length and relationship involvement. As can be seen in Table 3, relationship length was associated with confidence on the SHQ, and relationship involvement was associated with confidence on all four questionnaires.

Of course, because perceivers provided both confidence and involvement ratings, the observed relations between those variables may have been due to response biases. To evaluate this possibility, we asked if the extent to which targets reported being involved in the relationship predicted perceiver confidence. Target-reported involvement was a significant predictor of perceiver confidence on the SHQ, SLC, and APQ (βs = .36, .37, and .34, respectively, all *p*s < .05), suggesting that relationship involvement is associated with confidence for reasons above and beyond response biases. Overall, then, it appears that relationship length was somewhat associated with confidence, and relationship involvement was strongly associated with confidence.

Table 3
Predicting Confidence and Within-Dyad Accuracy From Relationship Length and Involvement (Study 1)

Measure	Beta weight: Relationship length	Beta weight: Relationship involvement
	Confidence	
SHQ	.32*	.30*
SLC	.00	.48**
SAQ	−.07	.28*
APQ	−.15	.49**
	Within-dyad accuracy	
SHQ	.24	−.07
SLC	.01	.19
SAQ	.00	−.07
APQ	−.06	.16

Note. SHQ = sexual history questionnaire; SLC = Self-Liking/Competence Scale; SAQ = Self-Attributes Questionnaire; APQ = Activity Preferences Questionnaire.
* $p < .05$. ** $p < .01$.

Accuracy. Next we tested the relation of relationship length and involvement to both within- and between-dyad accuracy. We began by regressing our within-dyad accuracy correlations on relationship length and involvement. The beta weights in Table 3 show that these analyses yielded no statistically significant relations.

We then conducted moderated multiple regressions to test whether between-dyad accuracy on each of our four questionnaires was moderated by relationship length or involvement. We regressed targets' self-ratings on (a) perceivers' predictions, (b) relationship length (or involvement), and (c) the interaction of (a) and (b). These regressions revealed that relationship length did not moderate between-dyad accuracy on any of our questionnaires ($|t|$s < 1.6, *p*s > .14). Relationship involvement did not moderate between-dyad accuracy on the SHQ or SLC (*t*s < 0.8, *ns*). However, involvement did moderate between-dyad accuracy on the APQ, $t(53) = 3.1$, $p < .003$ ($\beta = .44$), and it had a marginal moderating influence on SAQ between-dyad accuracy, $t(53) = 1.8$, $p < .09$ ($\beta = .21$), such that confidence increased with increases in involvement.

In short, although length and involvement were significantly related to confidence in five of eight tests, they predicted between-dyad accuracy in only one of eight tests (two of eight if we include the one marginal effect). This evidence that relationship length and involvement foster confidence despite being virtually unrelated to within- and between-dyad accuracy supports the idea that confidence depends on representational rich-

[5] Here and throughout the article, the *t* tests reported for regressions are the tests of the beta weights.

ness, whereas accuracy is largely unaffected by representational richness. We present more direct evidence for the role of representational richness in Study 3.

Were Perceivers Under- or Overconfident?

Dissociations of confidence and accuracy could occur among people who are generally overconfident, generally underconfident, or both under- and overconfident. We sought to examine the extent to which perceivers in the present study were under- or overconfident by comparing the average number of items on which they *actually* correctly predicted their partners' responses on each of our questionnaires with the average number of items on which they *thought* they correctly predicted their partners' responses. That is, we first gave perceivers 1 point each time one of their predictions exactly matched the partner's response and zero points each time a prediction was not exactly correct. We then compared this "actual hits" variable with perceivers' "estimated hits." The results indicated that perceivers were quite overconfident. Actual:estimated frequency ratios were 3.6:6.1 on the SHQ, 8.7:13.1 on the SLC, 4.5:7.8 on the SAQ, and 10.8:21.8 on the APQ (all within-participants $ts > 3.4$, $ps < .005$). Hence people thought that they accurately predicted their partners' responses far more often than they actually did, and such overconfidence was consistent across measures. In the defense of perceivers, although they were strikingly overconfident, their hit rates did exceed chance (44% vs. 20% on the SLC, 30% vs. 10% on the SAQ, and 29% vs. 14% on the APQ; all χ^2s$(1, N = 57) > 4.4$, $ps < .04$).[6]

In summary, the results of Study 1 suggest that confidence and accuracy are often dissociated in everyday person perception. We suggest that this dissociation occurs, in part, because confidence depends on representational richness, whereas accuracy does not. Our evidence that representational richness (as measured by relationship length and involvement) was more closely associated with confidence than accuracy is consistent with this view. Moreover, the results of the hits analyses suggest that people have a tendency to think that they know more than they actually do know about their partners.

Nevertheless, an alternative explanation may explain at least a portion of the data from Study 1. Specifically, confidence may have been (weakly) associated with relationship length because couples who did not feel especially familiar with one another broke up early in their relationships, resulting in relatively high levels of confidence among couples who stayed together. Such differential attrition could explain why couples in older relationships displayed relatively high levels of confidence. We attempted to rule out this possibility by using a longitudinal design in Study 2. In addition, to test the generality of our results, in Study 2 we examined a different type of relationship (same-sex roommates) and included some new measures.

Study 2: Confidence and Accuracy Among College Roommates

Method

Participants

Participants were 10 male and 30 female roommates residing in dormitories at the University of Texas at Austin. They ranged in age from 17 to 22, with a mean age of 18. When asked, 7.5% of participants reported that they were Black, 60% White, 10% Asian, 10% Hispanic, 5% Mexican American, 2.5% Puerto Rican, and 5% other. Roommates were unacquainted with one another prior to the beginning of the semester. Participants received $8 for their participation.

Procedure

Before the beginning of the semester, we placed 350 questionnaire packets in randomly selected mailboxes in each of 10 dormitories. The instructions on these packets urged participants not to discuss their responses with their roommate. Of the 350 questionnaire packets distributed, 146 (42%) were returned. Of these, 40 came from persons whose roommate did not return a packet, and 38 came from roommates who were already well acquainted; these participants were paid for completing the first packet but were not invited to complete the second. We contacted the remaining 68 roommates by telephone and invited them to complete a second questionnaire packet. Of these 68 persons, 40 returned to our laboratory and completed the second packet.

Participants completed identical questionnaire packets at the beginning of the semester (Time 1) and again 6 weeks later (Time 2). The questionnaires included a version of the SAQ that we had shortened by deleting the certainty and importance items and a shortened version of the SLC that included only the Self-Liking subscale. Roommates also rated themselves on 10 personality dimensions taken from the revised version of the Self-Attributes Questionnaires (SAQ-R; Pelham & Swann, 1989). The 10 items drawn from the SAQ-R were *sense of humor, extraversion, assertiveness, ambition, optimism, generosity, impatience, maturity, enthusiasm,* and *impulsiveness.* Finally, roommates rated the extent to which they enjoyed nine activities taken from the APQ, including *reading, watching TV, partying, seeing live music, watching sports, playing board games, cooking, working out,* and *cleaning.*

At both Time 1 and Time 2, participants rated themselves on each questionnaire and then attempted to predict how their roommate had responded. Also, they reported how confident they were (from 0–100%) that they had correctly predicted their roommates' responses on each questionnaire. To save time, participants made confidence ratings once for each questionnaire rather than for each item. At Time 2, perceivers estimated for each questionnaire how many of their roommates' responses they had predicted with perfect accuracy.

Roommates also answered, on 7-point scales, four questions about their involvement with their roommates. One of the questions asked how much time the respondent spent doing things with his or her roommate, a second focused on the amount of conversation that occurred between the roommates, a third tapped the extent to which participants shared problems or worries with their roommates, and the final question asked how much they liked their roommates. The index of relationship involvement consisted of the mean of these four items ($\alpha = .87$).

Computing Confidence Indices

The confidence indices consisted of a single item on the SAQ, SAQ-R, SLC, and APQ asking perceivers how confident they were that their predictions of their roommates' responses were accurate. As can be seen in Table 4, perceivers varied widely in the confidence they reported at both Time 1 and Time 2.

Computing Accuracy Indices

As in Study 1, we computed both within- and between-dyad accuracy correlations.

[6] The number of hits by chance on the SHQ were impossible to calculate because the questions were open-ended.

Within-dyad accuracy. Once again, our measures of within-dyad accuracy were intraclass correlations (Shrout & Fleiss, 1979) computed on each of our four questionnaires. These correlations suggested moderate levels of accuracy among our roommate pairs. As shown in Table 5, accuracy was much higher in some dyads than in others.[7]

Between-dyad accuracy. We computed average scores for targets and perceivers on each questionnaire at both Time 1 and Time 2. Internal consistencies for the SAQ-R and SLC were adequate at both points in time for both targets and perceivers (αs > .79). The SAQ showed borderline internal consistencies (αs = .59–.72), and internal consistency was low on the APQ (αs = .48–.51). The low internal consistency of the APQ suggests that it is not appropriate to average the items on that questionnaire. Thus, in Study 2 we examined between-dyad accuracy only on the SAQ, SAQ-R, and SLC.

At Time 1, between-dyad accuracy was $r(38) = .46$, $p < .004$ on the SAQ; $r(38) = .87$, $p < .001$ on the SAQ-R; and $r(38) = .02$, *ns* on the SLC. At Time 2, between-dyad accuracy was $r(38) = .15$, *ns* on the SAQ; $r(38) = .69$, $p < .001$ on the SAQ-R; and $r(38) = .39$, $p < .02$ on the SLC.

Table 5
Descriptive Statistics for Within-Dyad Accuracy Correlations (Study 2)

Measure	*M*	*SD*	Range
	Time 1		
SAQ	.37	.32	.00–.88
SAQ-R	.33	.27	.00–.86
SLC	.53	.35	.00–1.0
APQ	.44	.23	.00–.94
	Time 2		
SAQ	.48	.36	.00–.98
SAQ-R	.32	.26	.00–.87
SLC	.58	.36	.00–.99
APQ	.51	.28	.00–.97

Note. SAQ = Self-Attributes Questionnaire; SAQ-R = Self-Attributes Questionnaire–Revised; SLC = Self-Liking/Competence Scale; APQ = Activity Preference Questionnaire.

Results and Discussion

Relation of Confidence and Accuracy

We first examined the relation between the confidence and accuracy of people's impressions. We tested whether confidence was associated with both within- and between-dyad accuracy. In all cases, our four questionnaires were analyzed separately.

Confidence and within-dyad accuracy. We began by computing the correlation between perceiver confidence and our within-dyad accuracy correlations. At Time 1, accuracy and confidence were minimally related on the SAQ, $r(38) = -.16$; SAQ-R, $r(38) = -.10$; SLC, $r(38) = .06$; and APQ, $r(38) = -.12$. The same pattern held at Time 2 on the SAQ, $r(38) = -.07$; the SAQ-R, $r(38) = -.01$; and the APQ, $r(38) = -.07$. However, there was a relation between confidence and accuracy on the SLC, $r(38) = .40$, $p < .05$. These correlations replicate the results of Study 1, including the correlation between confidence and accuracy on the SLC.

Confidence and between-dyad accuracy. Next we examined whether the between-dyad accuracy correlation for each questionnaire was moderated by perceivers' confidence, using the moderated multiple regression technique we used in Study 1. These regressions revealed that confidence did not moderate between-dyad accuracy on any of our measures at either Time 1 or Time 2 ($|ts| < 1.3$, $ps > .19$). Taken together, Studies 1 and 2 suggest that confidence–accuracy dissociations occur across a wide array of personality judgments.

Table 4
Descriptive Statistics for Confidence Indices (Study 2)

Measure	*M*	*SD*	Range
	Time 1		
SAQ	63.2	23.5	2–100
SAQ-R	59.4	20.3	10–93
SLC	68.3	17.6	40–98
APQ	62.5	21.4	2–100
	Time 2		
SAQ	71.8	15.9	30–95
SAQ-R	65.1	16.6	30–100
SLC	72.1	16.3	35–100
APQ	72.1	16.9	35–100

Note. All numbers are percentages. SAQ = Self-Attributes Questionnaire; SAQ-R = Self-Attributes Questionnaire–Revised; SLC = Self-Liking/Competence Scale; APQ = Activity Preference Questionnaire.

Relation of Relationship Length and Involvement to Confidence and Accuracy

We expected that relationship length and involvement would produce increases in confidence but not accuracy.

Confidence. To test the relation between relationship length and confidence, we conducted a 2 (time: Time 1 vs. Time 2) × 4 (measure: SAQ vs. SAQ-R vs. SLC vs. APQ) multifactor repeated-measures analysis of variance (ANOVA). This analysis revealed the predicted main effect of time, $F(1, 39) = 7.5$, $p < .01$, such that perceivers' Time 1 predictions were made less confidently ($M = 63\%$) than their Time 2 predictions ($M = 70\%$). There was also a main effect of measure, $F(3, 117) = 8.1$, $p < .001$, such that the SLC predictions were made most confidently ($M = 70\%$), followed by SAQ and APQ predictions (Ms = 67%), followed by SAQ-R predictions ($M = 62\%$). There was no Time × Measure interaction, $F(3, 117) = 1.5$, $p = .21$, suggesting that the Time 1–Time 2 difference in confi-

[7] Using the pseudodyads technique (Corsini, 1956) to estimate stereotype accuracy, we discovered that at both Time 1 and Time 2, total accuracy was significantly higher than stereotype accuracy on the SAQ, SAQ-R, and APQ. On the SLC total accuracy was not greater than stereotype accuracy at either Time 1 or Time 2 (both $ps > .46$). At Time 1, estimates of stereotype accuracy were .25, .23, .50, and .19 for the SAQ, SAQ-R, SLC, and APQ, respectively. At Time 2 the respective estimates were .22, .20, .56, and .24. This suggests that accuracy correlations for the SAQ, SAQ-R, and APQ reflect true accuracy as well as stereotype accuracy, whereas accuracy correlations for the SLC primarily reflect stereotype accuracy.

dence did not vary as a function of which questionnaire was being examined. The change in confidence from Time 1 to Time 2 is presented in Table 6 separately for each questionnaire.

To test the relation between relationship involvement and confidence, we computed correlations between our involvement index (measured at Time 2) and Time 2 confidence on each of our measures. As can be seen in Table 6, these correlations were all statistically significant and of moderate magnitude. These correlations do not seem to be due to perceiver response biases, as indicated by follow-up analyses attempting to predict perceiver confidence from target-reported involvement. Targets' reports of involvement were significantly related to perceiver confidence on the SAQ, SAQ-R, and APQ (rs = .56, .37, and .47, respectively, $ps < .05$) and marginally related to perceiver confidence on the SLC ($r = .29$, $p < .07$).[8]

Accuracy. We tested the relation of relationship length and involvement to both within- and between-dyad accuracy. We began by submitting our within-dyad accuracy correlations to a 2 (time: Time 1 vs. Time 2) × 4 (measure: SAQ vs. SAQ-R vs. SLC vs. APQ) multifactor repeated-measures ANOVA. The effect of time on within-dyad accuracy was nonsignificant, $F(1, 39) = 2.5$, $p = .12$. (See Table 6 for the changes in within-dyad accuracy from Time 1 to Time 2 for each questionnaire.) Also, there was a significant effect of measure, $F(3, 117) = 2.9$, $p < .05$, such that within-dyad accuracy was highest on the SLC ($M = .56$), followed by the APQ ($M = .48$), the SAQ ($M = .43$), and the SAQ-R ($M = .33$). The Time × Measure interaction was not significant, $F(3, 117) = 1.2$, $p = .31$, suggesting that the Time 1–Time 2 difference in accuracy did not differ as a function of which measure was being examined. Overall, then, there was a very weak, nonsignificant trend toward increasing within-dyad accuracy, but this trend did not approach the magnitude of the increment in confidence over time.

Table 6
Predicting Confidence and Within-Dyad Accuracy From Relationship Length and Involvement (Study 2)

Measure	T1–T2 change	Correlation with T2 involvement
	Confidence	
SAQ	+9%	.63**
SAQ-R	+6%	.53**
SLC	+4%	.46**
APQ	+10%	.57**
	Profile accuracy	
SAQ	+.11	−.21
SAQ-R	−.01	.04
SLC	+.05	.11
APQ	+.07	−.25

Note. Multivariate analyses of variance revealed that the overall Time 1 (T1)-versus-Time 2 (T2) difference in confidence was significant, whereas the T1-versus-T2 difference in accuracy was not. The effect of time on confidence and accuracy did not differ as a function of which measure was being examined. SAQ = Self-Attributes Questionnaire; SAQ-R = Self-Attributes Questionnaire–Revised; SLC = Self-Liking/Competence Scale; APQ = Activity Preference Questionnaire.
** $p < .01$.

We tested whether within-dyad accuracy was related to relationship involvement (measured at Time 2) by computing correlations between our Time 2 measures of each. As can be seen in Table 6, there were no statistically significant relations between relationship involvement and within-dyad accuracy. In fact, there was a slight (nonsignificant) tendency for involvement to be negatively related to within-dyad accuracy on the SAQ and APQ.

We tested the moderating impact of relationship length on between-dyad accuracy by comparing between-dyad accuracy correlations at Time 1 and Time 2 using Fischer's r-to-z transformation.[9] These comparisons revealed a statistically significant decrease in between-dyad accuracy from Time 1 to Time 2 on the SAQ ($z = 2.5$, $p < .02$; rs = .46 and .15 at Time 1 and Time 2, respectively) and the SAQ-R ($z = 3.0$, $p < .003$; rs = .87 and .69 at Time 1 and Time 2, respectively). However, between-dyad accuracy did increase from Time 1 to Time 2 on the SLC ($z = 2.7$, $p < .007$, rs = .02 and .39 at Time 1 and Time 2, respectively). As noted above, we did not examine between-dyad accuracy on the APQ because of the low internal consistency of that measure. Next we used moderated multiple regressions to test whether between-dyad accuracy was moderated by relationship involvement. These regressions revealed that between-dyad accuracy was not significantly moderated by involvement on any of our four questionnaires ($|ts| < 1.7$, $ps > .11$). The largest t was for the SAQ ($t = -1.6$, $p = .11$) and suggested that, if anything, between-dyad accuracy decreased as relationship involvement increased.

Were Perceivers Under- or Overconfident?

As in Study 1, perceivers were quite overconfident. Actual:estimated frequency ratios were 4.15:6.15 on the SLC, 1.0:2.95 on the SAQ, 2.28:5.58 on the SAQ-R, and 2.5:5.63 on the APQ (all within-participants $ts > 4.0$, $ps < .001$). The results also resembled those from Study 1 in that perceivers' hit rates usually exceeded chance (42% vs. 20% on the SLC, 25% vs. 10% on the SAQ-R, and 27% vs. 14% on the APQ; all χ^2s$(1, N = 40) > 5.0$, $ps < .03$, except for the SAQ, 20% vs. 10%, $\chi^2(1, N = 40) = 1.1$, $p = .29$). The important point here, however, is that people thought that they accurately predicted their roommates' responses far more often than they actually did.

Taken together, the results of Studies 1 and 2 suggest that confidence–accuracy dissociations are substantial—confidence and accuracy were virtually unrelated on the majority of our measures. Study 2 revealed a stronger impact of acquaintanceship on confidence than did Study 1, probably because of the use of a longitudinal rather than a cross-sectional design. Finally, the fact that Study 2 showed an acquaintanceship effect on

[8] Because members of each roommate pair served as both target and perceiver in our analyses, our observations are not independent (Kenny & Judd, 1986). To assess the impact that this nonindependence might have had on our statistical analyses, we recomputed the effects of length and involvement on perceiver confidence with target confidence covaried out of perceiver confidence. Our effects were unchanged by this procedure, suggesting that nonindependence did not bias our analyses.

[9] Moderated multiple regression could not be used because relationship length was not a continuous variable in Study 2.

confidence using a longitudinal design argues against a "differential attrition" explanation of Study 1.

Study 3: The Mediational Role of Representational Richness

Although the results of Studies 1 and 2 are consistent with our proposal that representational richness leads to greater confidence in impressions, in both studies relationship length and involvement were merely proxies for representational richness. In Study 3 we measured representational richness in an attempt to provide more direct evidence of its mediational role. Specifically, we tested the hypotheses that relationship length and involvement foster rich representations and that rich representations foster confidence.

Method

Participants

We recruited 60 individuals involved in dating relationships by advertising in the University of Texas newspaper. Participants were paid $5. Participants ranged in age from 18 to 27 (M = 20), and the lengths of their relationships ranged from 2 weeks to 208 weeks (M = 50). As in Study 1, all participants were not living with their dating partner. When asked, 6% reported that they were Black, 58% White, 18% Asian, and 18% Hispanic. We dropped 1 participant from our analyses because he did not complete all the measures, another because she had ended her dating relationship 6 months prior to participating in our survey, and 3 more because our computer failed to save their data. This left 55 participants: 21 men and 34 women.

Procedure

We escorted each participant to a cubicle equipped with a computer. There, he or she completed one paper-and-pencil questionnaire and one computer-administered survey, with the two measures being counterbalanced across participants. The paper-and-pencil questionnaire contained several demographic questions, an item about relationship length, and the eight items composing our measure of relationship involvement (taken from Rusbult, 1980). The computer-administered survey was created using Superlab for Macintosh and consisted of having participants predict their dating partners' responses to the shortened versions of the APQ, SAQ-R, and SAQ and the full-length SHQ and SLC. The computer recorded the latency of each response; these latencies provided one measure of the richness of participants' impressions of their partners. Participants also completed one confidence item for each of the five scales on which they predicted their partners' responses. Each confidence item was answered on a 7-point scale, and the scale points were labeled in 10% increments ranging from 40% to 100%.

Finally, to gather more data relevant to the richness of participants' impressions, we gave participants 5 min to write an open-ended description of their partners' personalities. Two independent judges later rated these descriptions. Specifically, judges rated the amount of information conveyed in each description on a scale ranging from 1 (*minimal*) to 7 (*extensive*) as well as the degree of integration in the descriptions. Ratings of *integration* required familiarizing judges with the construct, which we defined as

> the number of connections the person has made among the pieces of knowledge he or she has about his/her partner. A person with an integrated representation of his/her partner can quickly think of information about the partner across a wide variety of domains due to a wealth of connections. Thus, you should rate a perceiver's open-ended description as integrated if it seems to present many different aspects of his/her partner rather than just one or two.

Interjudge agreement was .83 for the ratings of amount of information and .79 for the ratings of integration of information.

Results and Discussion

Relation of Relationship Length and Involvement to Confidence

To determine if our findings replicated the results of Studies 1 and 2, we conducted multiple regressions for each of our five measures, with confidence as the criterion variable and relationship length and relationship involvement as the predictor variables. As can be seen in Table 7, confidence was associated with length and involvement at above-chance levels. Similar to the results of Study 1 (which was also cross-sectional), involvement was a somewhat stronger predictor of confidence than was relationship length (four significant relations vs. one).

Relations of Relationship Length, Involvement, Representational Richness, and Confidence

To estimate the relations among length, involvement, representational richness, and confidence, we used the LISREL 8 software program (Jöreskog & Sörbom, 1996). Before conducting this analysis, we aggregated our data so that the number of parameters estimated for our model would be reasonable given our sample size. Although there is no consensus regarding the number of model parameters that can be estimated at a given sample size, we were guided by the notion that one's sample size should be at least five times greater than the number of parameters one estimates (Bentler & Chou, 1987).

Data aggregation. First we created two indices of representational richness. The first index of representational richness was based on judges' ratings. It consisted of the average of the two judges' ratings of the amount and integration of information in participants' open-ended descriptions of their partners (α = .93). The second index of richness was based on the accessibility of participants' responses. We began by deleting any response latencies that were greater than 2 standard deviations

Table 7
Predicting Confidence From Relationship Length and Involvement (Study 3)

	Beta weight	
Measure	Relationship length	Relationship involvement
APQ	.29*	.28*
SAQ-R	.15	.38**
SHQ	.13	.38*
SLC	−.16	.23
SAQ	.03	.37**

Note. APQ = Activity Preference Questionnaire; SAQ-R = Self-Attributes Questionnaire–Revised; SHQ = sexual history questionnaire; SLC = Self-Liking/Competence Scale; SAQ = Self-Attributes Questionnaire.
* $p < .05$. ** $p < .01$.

above the mean (Shoben, 1982). Next, we averaged participants' response latencies on each of our five questionnaires (αs = .79, .82, .76, .86, and .78 for the APQ, SAQ-R, SHQ, SLC, and SAQ, respectively) and then calculated the average of these five averages (α = .86). Finally, we created two indices of confidence by averaging the confidences on the SLC and SAQ-R and the confidences on the APQ, SHQ, and SAQ. This meant that each confidence index was based on an approximately equal number of items (i.e., SLC + SAQ-R = 30 items; APQ + SHQ + SAQ = 24 items). (As we show later, the particulars of the manner in which we constructed these confidence indices had little impact on the fit of our LISREL model.)

Specifying and testing the model. The model depicted in Figure 1 was based on treating judges' ratings and accessibility scores as indicators of the latent variable of representational richness. (Latent variables are unobservable variables that are presumed to "explain" the intercorrelations among a set of observed variables [Loehlin, 1992].)[10] We treated our two confidence indices as indicators of the latent variable of confidence; relationship length and relationship involvement were observed variables. We expected length and involvement to be positively related to representational richness and richness to be positively related to confidence.

We used the maximum likelihood estimation procedure to generate the standardized parameter estimates that are superimposed on the model in Figure 1.[11] To assess the fit of our model, we used chi-square, the goodness-of-fit index (GFI; Jöreskog & Sörbom, 1996), and the normed fit index (NFI; Bentler & Bonett, 1980). By each of these criteria, our model fit the observed data well: $\chi^2(8, N = 55) = 5.9$, $p = .66$, GFI = .97, and NFI = .93.[12] Furthermore, as can be seen in Figure 1, the predicted positive relations between length and representational richness, involvement and richness, and richness and confidence emerged.

Testing mediation. Following Baron and Kenny's (1986) reasoning, we assumed that if representational richness mediates length and involvement effects on confidence, then the addition of paths running directly from length and involvement to confidence should not improve the fit of our model. We accordingly created two additional models. One model included a direct path from length to confidence; the other included a direct path from involvement to confidence. We then compared these two models to our original model. Specifically, we subtracted the chi-square associated with each new model from the chi-square associated with our original model and tested the statistical significance of this new chi-square variable (with degrees of freedom equal to the difference between the degrees of freedom associated with the new and original models). We discovered that the addition of a direct path from length to confidence did not improve the fit of our model, nor did the addition of a direct path from involvement to confidence (both χ^2s < 1, *ns*).[13] These results support the conclusion that representational richness mediates the relations among length, involvement, and confidence.

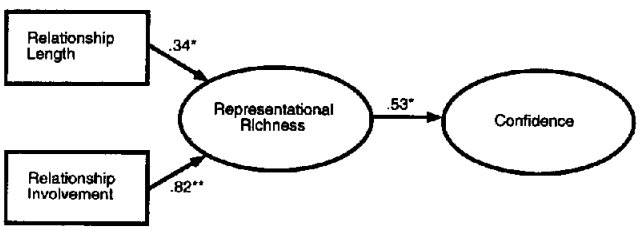

Figure 1. Model of the relations among relationship length, relationship involvement, representational richness, and confidence. Standardized maximum likelihood parameter estimates are superimposed on the model. *p < .05. **p < .01.

General Discussion

Our findings suggest that the confidence people have in their impressions of others is, at best, sporadically related to the accuracy of those impressions. Such confidence–accuracy dissociations seem to occur because people base the confidence of their impressions on cues that are largely unrelated to accuracy. Specifically, people base their confidence on the richness of their impressions. What they seem not to recognize is that the richness of impressions is not equivalent to the veracity of impressions. As a result, people in long-term, committed relationships may be quite bullish about the accuracy of beliefs that are no more accurate than the beliefs of people who have just initiated their relationships.

Our evidence that people base their confidence on the richness of their impressions fits with several themes in the cognitive literature. For example, this evidence is congenial to the notion that confidence is determined by the accessibility of information that seems relevant to a response rather than by direct access

[10] For example, general intelligence is a latent variable often used to explain the positive correlations that exist among a variety of measures of cognitive performance, with the assumption being that the measures are correlated because they all tap general intelligence. Latent variables account only for shared variance (and not error) among observed variables and thus are sometimes referred to as "error free." The observed variables whose intercorrelations are explained by the latent variable are called *indicators* of that latent variable.

[11] To set the scale of our latent variables (Jöreskog & Sörbom, 1996), the loading of judges' ratings on representational richness was set to 1.0, as was the loading of our first confidence index on the latent variable of confidence. Accessibility and our second confidence index had strong, statistically significant loadings on the prescribed latent variable: The standardized loading of our measure of accessibility on representational richness was −.52, and the standardized loading of our second confidence index on confidence was .97 (both ps < .01). Our model did not allow the residuals of any indicators to correlate. The correlation between length and involvement was not estimated; the observed correlation between those variables was quite small (r = .17, *ns*).

[12] Adequate fit of the model to the data is indicated by a nonsignificant chi-square and a GFI and NFI greater than .90 (Bollen, 1989). To examine whether the fit of our model depended on the particular manner in which we aggregated confidence scores, we aggregated confidence scores in several different ways and reran our model using each of these permutations. The fit of the model was always good according to all criteria, χ^2s(8, N = 55) < 10.98, ps > .14; GFIs and NFIs > .90.

[13] We conducted these same mediational tests using the models that were earlier used to examine the impact of aggregating confidence in different ways. We found that direct paths from length or involvement to confidence never improved the fit of our model (χ^2s < 1.3, *ns*) regardless of how our indicators of confidence were computed.

to an accurate response (Kelley & Lindsay, 1993; Koriat, 1993; Nelson & Narens, 1990). Our data also dovetail with Jacoby et al.'s (1994) suggestion that people's reliance on subjective indicators of knowing may be problematic because subjective indicators do not always correlate with objective indicators of knowledge (see also Wells & Murray, 1984). Finally, our evidence that participants overestimated the number of accurate predictions they made about their partners fits with the literature on overconfidence (e.g., Lichtenstein, Fischhoff, & Phillips, 1982). In fact, we even encountered evidence of overconfidence using a measure (frequency of "hits") that previous researchers have suggested is immune to overconfidence effects (Gigerenzer, 1991).

Advocates of the accuracy of the person-perception process (e.g., Funder, 1995) might be surprised by our evidence that accuracy rarely increased with length of relationship and that this occurred whether we computed accuracy within or between dyads. We believe that this finding fits with past work, however. For example, although the results of cross-sectional studies suggest that people in older relationships enjoy more accurate perceptions than people in newer relationships (Colvin & Funder, 1991; Funder & Colvin, 1988), longitudinal studies have failed to corroborate this conclusion consistently (Kenny, 1994). Conceivably, the apparent gains in accuracy in the cross-sectional studies may reflect a tendency for people to break off relationships characterized by inaccuracy, thus giving the appearance that people in older relationships have grown increasingly accurate despite being no more accurate than they were earlier on. Alternatively, the inherent nondiagnosticity of the small amounts of information available in new relationships may push accuracy below levels expected among unacquainted persons (Kenny, 1994); if accuracy later improves to the (quite modest) levels common among unacquainted persons, observers who did not examine accuracy at zero acquaintance may conclude erroneously that accuracy has risen to impressive levels. Clearly, the impact of acquaintanceship on accuracy demands more research, because so many attempts to describe a general pattern of results (e.g., Kenny, 1994) have subsequently been lost in a sea of contradictory data (e.g., Paulhus & Reynolds, 1995).

Our major focus here, however, was not on the absolute magnitude of accuracy but on the relation between confidence and accuracy. We believe that people's lack of insight into the accuracy of their beliefs could cause them to base important decisions on erroneous but confidently held beliefs. Consider the manner in which many people seem to be coping with the threat of AIDS. Recent research has suggested that at least some people have misunderstood former Surgeon General C. Everett Koop's injunction to "know your partner" to mean "learn about your partner's personality." They thus rely on their ability to form accurate impressions of personality to protect them from having sex with someone who is infected with HIV (e.g., Williams et al., 1992). This warped version of the know-your-partner strategy could lead people into trouble. For one thing, people's beliefs about potential partners may be inaccurate because of a tendency to use nondiagnostic cues to infer riskiness (e.g., the extent to which people seem familiar) and an inability to tell when others are lying about their sexual history (e.g., Swann, Silvera, & Proske, 1995). Paired with a tendency to base confidence on the richness of their beliefs, this tendency to form inaccurate impressions could cause people to become ever more confident in the safety of an unrelenting liar who is HIV positive.

Confidence–accuracy dissociations may also lead to disharmony in relationships. For example, difficulties could arise if a wife becomes highly confident that her husband is unsociable when in reality he suffers from performance anxiety. Similarly, a man's *lack* of confidence in his partner's professional aptitude could convince him to refrain from providing her with the encouragement that would inspire her to excel.

Let us add two caveats, however. First, the confidence–accuracy dissociations mentioned in the foregoing examples and those explored in our research all involve perceivers' inferences about the behavior of targets in multiple contexts, a relatively global form of accuracy. Often, however, perceivers have the much more modest goal of predicting the behavior of targets in their own presence only. Such circumscribed accuracy—which we did not measure—may be considerably easier to attain (e.g., Athay & Darley, 1981; Swann, 1984) and may be strongly related to confidence.

Second, even if it turns out that confidence–accuracy dissociations *are* quite pervasive, they may sometimes be relatively benign. After all, in most situations most of the time, normative pressures and interpersonal loyalties ensure that people do not engage in behaviors that might place their partners at grave physical or psychological risk. In addition, even if a wayward target does engage in a behavior that would cause his or her partner some distress, the partner may never learn of the transgression and may therefore never experience distress. Finally, the potential costs of lack of insight into the accuracy of beliefs may be balanced by a tendency for confidence to be inherently beneficial. Confidence may, for example, engender a feeling of comfort because we feel that we know what to expect from the people around us. From this vantage point, the key to interpersonal harmony may be not that confidence grows with accuracy but simply that confidence grows.

References

Albright, L. (1990). [A longitudinal study of consensus and accuracy in interpersonal perception.] Unpublished raw data.

Andersen, S. M. (1984). Self-knowledge and social inference: II. The diagnosticity of cognitive/affective and behavioral data. *Journal of Personality and Social Psychology, 55,* 294–307.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51,* 1173–1182.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588–600.

Bentler, P. M., & Chou, C. (1987). Practical issues in structural equations modeling. *Sociological Methods and Research, 16,* 78–117.

Berger, I. E., & Mitchell, A. A. (1989). The effect of advertising on attitude accessibility, attitude confidence, and the attitude–behavior relationship. *Journal of Consumer Research, 16,* 269–279.

Bernieri, F. J. (1991). Interpersonal sensitivity in teaching interactions. *Personality and Social Psychology Bulletin, 17,* 98–103.

Bernieri, F. J., Zuckerman, M., Koestner, R., & Rosenthal, R. (1994). Measuring person perception accuracy: Another look at self–other agreement. *Personality and Social Psychology Bulletin, 20,* 367–378.

Berscheid, E., Graziano, W., Monson, T., & Dermer, M. (1976). Out-

come dependency: Attention, attribution, and attraction. *Journal of Personality and Social Psychology, 34,* 978–989.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.

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.

Corsini, R. J. (1956). Understanding and similarity in marriage. *Journal of Abnormal and Social Psychology, 52,* 327–332.

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

Dunning, D., Griffin, D. W., Milojkovic, J. D., & Ross, L. (1990). The overconfidence effect in social prediction. *Journal of Personality and Social Psychology, 58,* 568–581.

Fazio, R. H., & Zanna, M. P. (1978). On the predictive validity of attitudes: The roles of direct experience and confidence. *Journal of Personality, 46,* 228–243.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102,* 652–671.

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.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases." In W. Stroche & M. Hewstone (Eds.), *European review of social psychology* (Vol. 2, pp. 83–115). Chichester, England: Wiley.

Gill, M. J., Swann, W. B., Jr., & Silvera, D. H. (1997). *On the genesis of confidence.* Manuscript submitted for publication.

Jacoby, L. L., Bjork, R. A., & Kelley, C. M. (1994). Illusions of comprehension, competence, and remembering. In D. Druckman & R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance* (pp. 57–80). Washington, DC: National Academy Press.

Jöreskog, K., & Sörbom, D. (1996). LISREL 8: User's reference guide. Chicago: Scientific Software International.

Kelley, C. M., & Lindsay, D. S. (1993). Remembering mistaken for knowing: Ease of retrieval as a basis for confidence in answers to general knowledge questions. *Journal of Memory and Language, 32,* 1–24.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* New York: Guilford Press.

Kenny, D. A., & Judd, C. M. (1986). Consequences of violating the independence assumption in the analysis of variance. *Psychological Bulletin, 99,* 422–431.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100,* 609–639.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 306–334). New York: Cambridge University Press.

Loehlin, J. C. (1992). *Latent variable models* (2nd ed.). Hillsdale, NJ: Erlbaum.

Metzler, C. W., Noell, J., & Biglan, A. (1992). The validation of a construct of high-risk sexual behavior in heterosexual adolescents. *Journal of Adolescent Research, 7,* 233–249.

Murray, S. L., & Holmes, J. G. (1993). Seeing virtues in faults: Negativity and the transformation of interpersonal narratives in close relationships. *Journal of Personality and Social Psychology, 65,* 707–722.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 125–173). San Diego, CA: Academic Press.

Oskamp, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology, 29,* 261–265.

Paulhus, D. L., & Reynolds, S. (1995). Enhancing target variance in personality impressions: Highlighting the person in person perception. *Journal of Personality and Social Psychology, 69,* 1233–1242.

Paunonen, S. V., & Jackson, D. N. (1985). Ideographic measurement strategies for personality and prediction: Some unredeemed promissory notes. *Psychological Review, 92,* 486–511.

Pelham, B. W., & Swann, W. B., Jr. (1989). From self-conceptions to self-worth: On the sources and structure of global self-esteem. *Journal of Personality and Social Psychology, 57,* 672–680.

Pieters, R., & Verplanken, B. (1995). Intention–behavior consistency: Effects of consideration set size, involvement, and need for cognition. *European Journal of Social Psychology, 25,* 531–543.

Rusbult, C. E. (1980). Commitment and satisfaction in romantic associations: A test of the investment model. *Journal of Experimental Social Psychology, 16,* 172–186.

Ryback, D. (1967). Confidence and accuracy as a function of experience in judgment making in the absence of systematic feedback. *Perceptual & Motor Skills, 24,* 331–334.

Saunders, D. R. (1956). Moderator variables in prediction. *Educational and Psychological Measurement, 16,* 209–222.

Sherman, J. W., & Klein, S. B. (1994). Development and representation of personality impressions. *Journal of Personality and Social Psychology, 67,* 972–983.

Shoben, E. J. (1982). Semantic and lexical decisions. In C. Richard Puff (Ed.), *Handbook of research methods in human memory and cognition* (pp. 287–314). New York: Academic Press.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–434.

Smith, E. E., Adams, N. A., & Schorr, D. (1978). Fact retrieval and the paradox of interference. *Cognitive Psychology, 10,* 438–464.

Snodgrass, S. E., & Rosenthal, R. (1985). Interpersonal sensitivity and skills in decoding nonverbal channels: The value of face value. *Basic and Applied Social Psychology, 6,* 243–255.

Surra, C. A., & Longstreth, M. (1990). Similarity of outcomes, interdependence, and conflict in dating relationships. *Journal of Personality and Social Psychology, 59,* 501–516.

Swann, W. B., Jr. (1984). Quest for accuracy in person perception: A matter of pragmatics. *Psychological Review, 91,* 457–477.

Swann, W. B., Jr., & Ely, R. J. (1984). A battle of wills: Self-verification versus behavioral confirmation. *Journal of Personality and Social Psychology, 46,* 1287–1302.

Swann, W. B., Jr., Silvera, D. H., & Proske, C. U. (1995). On "knowing your partner": Dangerous illusions in the age of AIDS? *Personal Relationships, 2,* 173–186.

Tafarodi, R. W., & Swann, W. B., Jr. (1995). Self-liking and self-competence as dimensions of global self-esteem: Initial validation of a measure. *Journal of Personality Assessment, 65,* 322–342.

Wells, G. L., & Murray, D. M. (1984). Eyewitness confidence. In G. Wells & E. Loftus (Eds.), *Eyewitness testimony: Psychological perspectives* (pp. 155–170). New York: Cambridge University Press.

Williams, S. S., Kimble, D. L., Covell, N. H., Weiss, L. H., Newton, K. J., Fisher, J. D., & Fisher, W. A. (1992). College students use implicit personality theory instead of safer sex. *Journal of Applied Social Psychology, 22,* 921–933.

Received February 12, 1997
Revision received May 12, 1997
Accepted May 12, 1997 ■