Personality and Social Psychology Review
2006, Vol. 10, No. 4, 282–294

Componential Analysis of Interpersonal Perception Data

David A. Kenny
Tessa V. West
Department of Psychology
University of Connecticut

Thomas E. Malloy
Department of Psychology
Rhode Island College

Linda Albright
Department of Psychology
Westfield State College

We examine the advantages and disadvantages of 2 types of analyses used in interpersonal perception studies: componential and noncomponential. Componential analysis of interpersonal perception data (Kenny, 1994) partitions a judgment into components and then estimates the variances of and the correlations between these components. A noncomponential analysis uses raw scores to analyze interpersonal perception data. Three different research areas are investigated: consensus of perceptions across social contexts, reciprocity of attraction, and individual differences in self-enhancement. Finally, we consider criticisms of componential analysis. We conclude that interpersonal perception data necessarily have components (e.g., perceiver, target, measure, and their interactions), and that the researcher needs to develop a model that best captures the researcher's questions.

Interpersonal perception, broadly defined as "the study of the beliefs that people have about others" (Kenny, 1994, pp. vii), has long been a topic of considerable theoretical and empirical interest. The basic data in interpersonal perception minimally contain three key elements: perceivers, targets, and measures. In a typical study, a set of perceivers rate a set of targets on several measures. As an example, Israel (1958) had 29 nurses who served as both perceivers and targets who judged each other on orderliness, leadership ability, appearance, and intelligence.

Experimental research in interpersonal perception is straightforward. The researcher has an independent variable and the levels of the independent variable refer to different targets, perceivers, or measures. For instance, in a study of stereotyping, the independent variable might be race of the target (White or Black).[1] For example, in the hypothetical study of race, the researcher might average across measures and targets, separately for Black and White targets, and compute a paired *t* test to determine if there is a race effect.

However, not all the questions in interpersonal perception can be answered by experimental[2] methods. Many of the questions in interpersonal perception involve computing relationships between two different interpersonal perception measures. Consider the following examples:

- Meta-accuracy involves computing a relationship between how perceivers think others view them with how the others actually view them.
- Assumed similarity involves computing a relationship between perceptions of others and self-perceptions.
- Consensus involves computing the correlation between judgments made by two or more perceivers of the same target.

The concern of this article is not the design and analysis of experiments in interpersonal perception, but

We thank David Funder, Lee Jussim, and Marina Julian who provided comments on a prior draft of this article.

Correspondence should be addressed to David A. Kenny, Department of Psychology, University of Connecticut, Storrs, CT 06269–1020. E-mail: kenny@uconn.edu

[1]There is the complication of unit of analysis; most researchers use perceiver. As we discuss later, target and sometimes measure can be treated as random factors, which is inconsistent with using perceiver as the unit.

[2]We use the term *experimental* in the most general sense and not in the sense of requiring manipulation of the independent variable. For instance, a study of gender differences in leniency in the perception of others would be experimental.

rather studies that correlate two sets of interpersonal perceptions.

To answer questions when two sets of perceptions are correlated, we can adopt one of two fundamentally different orientations. Consider the question of assumed similarity: Do people see others as similar to themselves? Most researchers would simply correlate trait judgments of others with trait judgments of self. Other researchers would treat measures as being made up of components. If a componential strategy is used, there is no single answer to the assumed similarity question, but several answers. We refer to the former approach as *noncomponential* and to the latter approach as *componential*.

As an illustration of these two approaches, consider the following example. Table 1 presents the ratings made by two different perceivers of five different targets on a single measure. We also have a criterion measure or the "truth." The interest is in individual differences in accuracy: How accurate is each of the perceivers in his or her assessments? An obvious noncomponential measure of a judge's ability is the sum of the absolute discrepancies. Thus, for Perceiver A, we would compute the sum of the absolute differences between A's judgment and the truth. Using this measure, both perceivers are equally good or bad, both having a score of 6.00. Thus, from a noncomponential analysis, we would conclude that the two perceivers were equally accurate.

However, the componential approach views the judgments and the criterion as containing components: one component that refers to targets in general (the mean) and another that refers to each target in particular (the deviation from the mean). We can assess accuracy for each component, and therefore there are two different ways to assess accuracy or inaccuracy. A perceiver can be wrong about the level or average value of the targets on the measure, and a perceiver can be wrong about the relative standings of targets on the measure. The former is called *elevation accuracy* and the latter *differential accuracy* (Cronbach, 1955). Given that the mean of the criterion is 4.0, Perceiver A's mean is 5.2, and Perceiver B's mean is 4.0, Perceiver B is better in terms of elevation accuracy. Differential accuracy can be measured by correlating a perceiver's judgments with the truth. Perceiver A's correlation with the truth is .986, whereas Perceiver B's correlation with the truth is –.224. Thus, it turns out that Perceiver A is a better judge in terms of differential accuracy (i.e., the correlations), whereas Perceiver B is a better judge in terms of elevation accuracy (i.e., the means). A componential analysis yields a more complicated pattern of results: In one way A is the better judge and in another way B is the better judge. As this simple example illustrates, noncomponential and componential analyses can give very different results.

The purpose of this article is to elaborate on the benefits and drawbacks of componential and noncomponential analysis of interpersonal perception data. We begin by explaining what componential analysis is, and then, using three case studies, we illustrate empirical differences between componential and noncomponential analyses. In the last section of the article, we discuss criticisms that have been leveled at componential analysis.

Table 1. *Hypothetical Judgments by Two Perceivers of Five Targets.*

Target	A's Judgment	B's Judgment	Truth
1	5	3	4
2	4	5	3
3	3	4	2
4	6	4	5
5	8	4	6

What Is Componential Analysis?

A major difficulty in componential analysis is the understanding of what componential analysis exactly is. No doubt some of this confusion stems from the bewildering array of terms that Cronbach (1955) gave to his four types of accuracy: elevation, differential elevation, stereotype accuracy, and differential accuracy. Moreover, Cronbach was not very explicit that those types of accuracy involved relationships between components.

As we stated in the beginning of the article, the basic data in interpersonal perception contain three key elements: perceivers, targets, and measures. In a hypothetical study, we gather data from all perceivers judging all targets on all measures. The design can be viewed as a three-way analysis of variance (ANOVA) of Perceiver × Target × Measure. We describe three different types of partitioning the variance of interpersonal perception data: Cronbach, Social Relations Model, and Judd and Park.

Cronbach Partitioning

Cronbach (1955) considered the judgments of a single perceiver of a set of targets on a set of measures. His data structure is a two-way ANOVA with judgments decomposed into target and measure:

$$\text{Judgment} = \text{Mean} + \text{Target} + \text{Measure} + (\text{Target} \times \text{Measure})$$

Cronbach (1955) was primarily interested in accuracy, and, therefore, a criterion measure is also measured. For instance, there might be self-ratings by the targets on each of the measures. The criterion measure can also be decomposed into the four ANOVA components:

Criterion = Mean' + Target' + Measure' + (Target × Measure)'

where the prime symbol is used to denote that the component refers to the criterion measure. Accuracy is then the association between components (see Figure 7.1 in Kenny, 1994). For example, differential accuracy is the association between the two Target × Measure interactions. The Cronbach approach is idiographic because the four accuracies (the associations between the four corresponding components) are measured for each perceiver.

Social Relations Model Partitioning

Here, we start with the same three-way data structure of Perceiver × Target × Measure. Within the Social Relations Model (SRM; Kenny & Albright, 1987), we consider for each measure a two-way structure of Judge × Target. These data are decomposed in a two-way ANOVA model:

Judgment = Mean + Perceiver + Target + (Perceiver × Target)

In traditional SRM parlance, Perceiver is called Actor, Target is called Partner, and Perceiver × Target is called Relationship.

Judd and Park (1993) Partitioning

Judd and Park's (1993) full accuracy design begins with three-way data structure of Perceiver × Target × Measure. They then consider an ANOVA decomposition of the data involving eight components: one mean, three main effects, and four interactions. A key feature of their approach is the selection of perceivers, targets, and measures. Specifically, the perceivers are members of two different groups (e.g., men and women), as are the targets. Half of the measures are stereotypical of members of one group (e.g., aggressive for men), and the other half are stereotypical of members of the other group (e.g., compassionate for women). Additionally, half of the measures are positively valenced (e.g., compassionate) and half are negatively valenced (e.g., aggressive).

Methodological Issues for Componential Analysis

For all three types of componential analysis that we have described, we see that the components are derived from an ANOVA decomposition of the data. Basic questions in interpersonal perception involve computing measures of association between these ANOVA components. There are some technical but important issues in componential analysis: whether components should be considered fixed or random, the particular measure of the association between components, and the nonindependence of components.

We begin by addressing the issue of fixed versus random components. In an ANOVA model, the factors can be considered as either fixed or random. Consider the two-way decomposition of Perceiver × Target for judgments of a single measure (e.g., Extroversion). One set of *n* perceivers judge *m* different targets. We then compute a two-way ANOVA and obtain mean squares due to perceiver or MS_P, mean squares due to target or MS_T, and mean squares due to the perceiver by target interaction or MS_{PxT}. If we consider targets as fixed, the variance due to targets is MS_T/n. If however, we consider targets as random, the ANOVA estimate of target variance is $(MS_T - MS_{PXT})/n$. Debate about whether components should be fixed or random is a long-standing one (Norman, 1967). Whereas treating components as random allows for greater generality, it also brings with it greater computational complexity. In the 21st century, the study of interpersonal perception deserves that complexity.

A second issue in componential analysis is determining how to measure the association between components. Consider the question of self-other agreement. In a componential analysis, we would compute the target effect (how a person is generally seen by others) and relate that component to how the person sees him- or herself. One can relate the two scores by computing a correlation coefficient between components, computing a difference score (see Judd & Park, 1993), or computing the absolute difference in the scores. The researcher needs to decide how to relate the components of interpersonal perception based on theoretical and methodological considerations.

Finally, in some interpersonal perception studies, the same people are both the perceivers and targets. For instance, in a round-robin design, a group of people judge one another. Such judgments result in nonindependent data. For example, perceiver A's judgment of target B might well be correlated with perceiver B's judgment of target A. The nonindependence in the data must be modeled.

We use the case study method to illustrate the differences between noncomponential and componential analyses for reciprocity of attraction (Kenny & Nasby, 1980), cross-context consensus (Malloy, Albright, Kenny, Agatstein, & Winquist, 1997), and self-enhancement (Kwan, John, Kenny, Bond, & Robins, 2004). These examples are quite diverse in terms of findings and topics. For all three questions, we use SRM partitioning of data because we are much more familiar with this approach. We believe that the same points could be illustrated with the other types of componential analysis.

Reciprocity of Attraction

Reciprocity has long been believed to be a fundamental principle of interpersonal attraction. Somewhat surprisingly, however, the social psychological literature contains comparatively few direct assessments of the extent to which reciprocal attraction actually characterizes interpersonal relationships. Even more surprising, these assessments indicated weak or even zero reciprocity. For instance, Newcomb (1961), in his classic study of the acquaintance process, obtained attraction ratings repeatedly from two groups of 17 college residence-hall students over a 15-week period. Although his results indicated that actual reciprocity became more pronounced as acquaintance increased (i.e., over time), even at the end of 15 weeks the degree of actual reciprocity had reached only modest proportions ($r_s \approx .345$, calculated from the distribution of differences in ranks of attraction assigned by members of all pairs to each other). Kenny and Nasby (1980) questioned whether reciprocity may merely represent another addition to the already lengthy parade of the attributional and cognitive biases that occur in social interactions. Indeed, specific demonstrations of a reciprocity bias have emerged from several experimental procedures, including estimates of subjective probability (Desoto & Kuethe, 1959), paired-associates learning (Zajonc & Burnstein, 1965), recognition memory (Sherman & Wolosin, 1973), and metaperception of friendship (Kenny, Bond, Mohr, & Horn, 1996).

Kenny and Nasby (1980) argued that a componential analysis can be used to explain the low reciprocity correlations. A liking judgment represents four different components:

$$X_{ijk} = m_k + a_{ik} + b_{jk} + g_{ijk} \qquad (1)$$

$$X_{jik} = m_k + a_{jk} + b_{ik} + g_{jik} \qquad (2)$$

where X_{ijk} is person i's attraction toward person j in group k, and X_{jik} is j's attraction toward i. The connections between the two measures are discussed later. The first term in Equation 1, m_k, is the grand mean and measures the average level of attraction across all raters and targets in group k. The second term, a_{ik}, is the perceiver effect for person i in group k and represents how much person i tends to like others in general. The third term, b_{jk}, is the target effect of person j in group k and represents how much person j is liked by others in general. The fourth term, g_{ijk}, is the Perceiver × Target interaction or relationship effect, which measures the degree to which perceiver i particularly likes the target j after subtracting out the perceiver effect for i and target effect for j.

There are three sources of variance or covariance that produce the reciprocity correlation, which are represented by the lines connecting person i's liking of j and j's liking of i in Equations 1 and 2. The line connecting the m_k terms represents the group norm for liking. If people are in groups in which individuals like one another, they would tend to like each other, and if people are in groups in which individuals did not like one another, they would tend to dislike each other. The two diagonal lines connect the individual differences in giving liking with the difference in receiving liking. There are two such connections, because there are two persons. This is the generalized correlation between giving and receiving liking and is denoted as r_1. The final connection between Equations 1 and 2 is between the interactions of i toward j and of j toward i. This correlation, then, measures the dyadic reciprocity and is denoted as r_2.

The data that we reanalyze were gathered by Curry and Emerson (1970) and include six eight-person groups who reported their attraction toward each other at five times, the first being one week after the first meeting. We consider a somewhat different question from Kenny and Nasby (1980). They treated the five time points as independent replications. We averaged the five time points, so we have the average liking of each perceiver of the target. Moreover, Kenny and Nasby subtracted off the group means and we did not.

What is the reciprocity correlation? If we ignore components and just compute the correlation between pairs of scores,[3] we find that the correlation is .318. However, that correlation combines all the components together. As we shall see, this single correlation contains within it three correlations. Unless we do a componential analysis, we do not know if the source of this modest correlation is the group, the individual, or the relationship. We can remove group effects by computing the correlation within groups. When we do so, the correlation reduces to .185 and is no longer statistically significant. The fact that the correlation gets smaller when group is controlled for suggests that groups differ from one another in how much they like one another (i.e., there must be group variance).

Table 2 presents the partitioning of the variance of the attraction measure from the Curry and Emerson (1970) pooled across the six clusters. We see that there is variance in all four components. However, the largest source of variance is at the relationship level, a result common in studies of attraction (Kenny, 1994). Therefore, how much one person likes another is idiosyncratic.

In terms of reciprocity, we find an interesting pattern. The reciprocity correlation at the individual level

[3] Because we have dyadic data with indistinguishable members, we can use the pairwise or double-entry method (Griffin & Gonzalez, 1995) to compute the correlation. We take each dyad and enter the score twice, once with person 1 as the X variable and once with person 2 as the X variable.

Table 2. *Variance Partitioning and Correlations for the Curry and Emerson (1970) Study.*

Relative Variances[a]	
Perceiver	.159
Target	.212
Relationship	.506
Group	.123
Correlations	
Naive	
Raw	.318
Raw With Group Removed	.185
SRM	
Generalized	–.210
Dyadic	.554

Note: SRM = social relations model.
[a]Proportion of total.

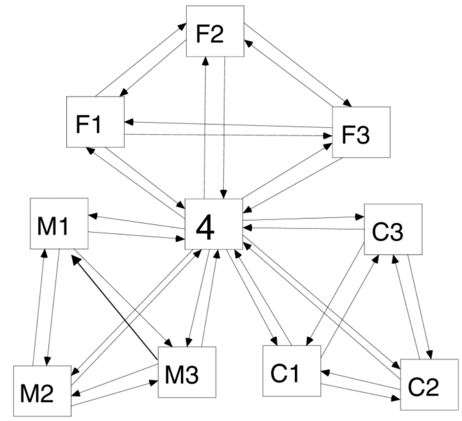

Figure 1. The key person design (F: Friend, M: Family, C: Coworker).

is negative, being –.210. Thus, it means that people who like others are not liked by others. This negative correlation is not, however, significantly different from zero. The reciprocity correlation at the dyadic level is .554 and is statistically significant. If one person especially likes another person (e.g., more than he or she generally likes others and more than the other is generally liked), that liking is reciprocated.

What have we learned from the componential analysis? First, we have learned that there is quite a bit of reciprocity that primarily occurs at the dyadic level. However, there is not generalized reciprocity. Moreover, a norm of liking and disliking raises the reciprocity correlation. Only by a componential analysis were we able to discover reciprocity. Without it, we would have found the same weak levels of reciprocity that the noncomponential literature has typically found.

We have seen that a componential analysis gives us a very detailed picture of reciprocity. We have found that reciprocity is largely due to relational, not individual, processes. The next case study explores the extent to which there is consensus on how a target individual is perceived within and between distinct social contexts.

Cross-Context Consensus

For our second case study, we reanalyze and reexamine the results of the Malloy et al. (1997) study. One reason we chose this study is because Malloy et al. argued that a noncomponential analysis of their data would be misleading. They presented extensive componential results, but did not feature noncomponential findings. We reanalyze their data set using both noncomponential and componential analyses and explain both theoretically and statistically why they differ.

Malloy et al. (1997) were interested in the degree to which there is consensus in how people are viewed in three distinct contexts: leisure, home, and work. Each participant rated and was rated by three friends (Leisure), three family members (Home), and three coworkers (Work). In this study, there were 31 participants and 279 (i.e., 31 × 9) informants. Thus, each participant was a member of three separate round-robin data structures. This design is presented in Figure 1. These three groups were nonoverlapping in that, with the exception of the participant, no member from one group was acquainted with any member from another group. Within each group of four, each person rated every member in the group and the participant (person 4), who was a member of all three groups.

Each of the four perceivers rated each target on five measures designed to indicate the Big Five factors (John, 1990): shy–outgoing (Extroversion), rude–courteous (Agreeableness), late–on time (Conscientiousness), calm–anxious (Emotional Stability), and unintelligent–intelligent (Culture). The central question that Malloy et al. (1997) addressed is: To what extent were participants seen the same way across different contexts?

We begin with the noncomponential approach to the question posed by Malloy et al. (1997). This strategy (e.g., Funder, Kolar, & Blackman, 1995) is to average the rating of the three informants' judgments within a context, and then correlate that average across contexts. We computed those correlations for each of the Big Five for the three pairs of contexts. We then averaged across the Big Five, and the resulting noncomponential correlations are presented in Table 3. We see that correlations are moderate in size, ranging from .26 to .41.[4]

Malloy et al. (1997) pointed out that these noncomponential correlations are potentially misleading. The authors argued that the SRM provides a much more accurate method of analysis by estimating vari-

[4]These results differ somewhat from those reported in Malloy et al. (1997) largely because of a coding error.

Table 3. *Cross-Context Correlations*[a].

	Work–Home	Work–Leisure	Home–Leisure
Noncomponential	.29	.41	.26
SRM Effects			
Partner[b]	.13	.33	.25
Mean[c]	.00	.25	.26
Mean (Adjusted)[d]	.07	–.04	.06

Note: SRM = social relations model.
[a]All correlations are the simple averages of five correlations. [b]Mean rating by all four people within the context. [c]Mean rating the three informants, excluding the participant as both a perceiver and a target. [d]Means computed without person 4's data.

ance due to perceiver, target, and relationship-error, than does a noncomponential analysis. In the SRM, the rating of perceiver i of target j in group k in context q, or X_{ijkq}, would be assumed to equal

$$m_{kq} + a_{ikq} + b_{jkq} + g_{ijkq}$$

where m_{kq} is the mean rating for group k in context q, a_{ikq} is the perceiver effect (how perceiver i in group k in context q rates others on average), b_{jkq} is the target effect (how target j is seen by others in group k in context q on average), and g_{ijkq} is the relationship effect (how perceiver i views target j in group k in context q).

To test the hypothesis that individuals are judged similarly across different contexts, we need to measure the SRM term b_{4kq}, the target effect for person 4 in the group (target person or participant who is in all three groups). Noncomponential and componential approaches estimate b_{4kq} in very different ways. In a noncomponential analysis, we average the three ratings of target 4. Within the SRM, that average can be shown to equal the following sum of components:

$$b_{4kq} + m_{kq} + \Sigma a_{ikq}/3 + \Sigma g_{i4kq}/3 \qquad (3)$$

The latter two terms can be treated as random error, in that they are not likely to be correlated across contexts. It is relatively clear that the relationship component or g would not correlate across contexts. The term $\Sigma a_{ikq}/3$ is the average perceiver effect within a context. Although it may seem plausible that individuals might pick friends, coworkers, and family members who have similar views of others, we can allocate that effect to the group effect.

The key theoretical question is how stable b_{4kq} is across contexts. For example, if a participant is generally seen by family members as intelligent, is that participant also seen as intelligent by coworkers and friends? When we examine the noncomponential estimate of b_{4kq} in Formula 3, we see that it contains m_{kq}, or the mean of how people in general are seen in that context. Because these group means might be correlated across contexts, there is a potential confound. Malloy et al. (1997) marshaled evidence (e.g., Snyder & Ickes, 1985) that this might well be the case. In particular, individuals may choose contexts, especially leisure and work, that are consistent with their family environment. Thus, the group effect is a potentially confounding variable for the noncomponential approach. If the cross-context correlation were due to the group effect or m_{kq}, it would be much less interesting, because it would indicate that two perceivers would agree in their rating of two different persons across the two contexts. For example, a family member and a coworker might agree that Joe is extroverted, but they may also believe that Joe's father and Joe's boss are also extroverted. The theoretical interest in terms of cross-situational validity of personality is whether two perceivers would agree in their rating of the same person across two contexts. Thus, the noncomponential correlation across contexts could be nonzero for two very different reasons: the partner effect, or b_{4kq}, and the group effect or m_{kq}.

It is rather complicated to compute the componential estimate of the partner effect, or b_{4kq}. Kenny (1994, p. 236) presented the following formula for its estimation:

$$b_i = \frac{(n-1)^2}{n(n-2)} M_{.i} + \frac{(n-1)}{n(n-2)} M_{i.} - \frac{(n-1)}{(n-2)} M_{..}$$

where n is the group size, $M_{.i}$ is the mean rating when person i is the target, $M_{i.}$ is the mean rating when person i is the perceiver, and $M_{..}$ is the mean across all judgments in the group.[5] Applying this formula when n is 4 for the estimate of the partner effect of person 4 in group k and context q yields

$$1.125M_{.4kq} + 0.375M_{4..kq} - 1.500M_{..kq} \qquad (4)$$

where $M_{.4k}$ is the mean of the three ratings received by person 4 in group k and context q, $M_{4.kq}$ is the mean rating made by person 4 of the others in group k in context q, and $M_{..kq}$ is the rating of the 12 judgments (all 4

[5]Note that the formula is not the mean rating of the target minus the grand mean, as might be naively thought. That formula would be biased by the person's actor effect. The complications in Equation 4 are due to the "missing data" for self-ratings.

perceivers' ratings of 3 targets) in group k in context q. Malloy et al. (1997) used Formula 4 to estimate b_{4kq}. Table 3 presents correlations of these target effects (again averaged over the Big Five) between pairs of contexts. They parallel the noncomponential correlations but are weaker, ranging from .13 to .33. Malloy et al. suggested that they are lower than the noncomponential correlations because of the confounding effect of the mean:

> It is likely that individuals seek out and select into groups that are similar to other groups to which they belong … Specifically, selection into similar social groups may result in correlated group means on a variable that bias (i.e., inflate) estimates of agreement across them. (p. 392)

In addition to controlling for the mean to obtain a more accurate target effect, componential analysis is, in principle, a more accurate analysis than raw score analysis because a purer measure of the target effect is obtained. The effects of perceiver and mean have been removed. If we take Formula 4 and substitute in the SRM components, the resulting estimated target effect (i.e., expected value) for person 4 in group k and context q equals

$$b_{4kq} - \Sigma_i b_{ikq}/4 - \Sigma\Sigma 0.125 g_{ijkq} + 0.250\ \Sigma g_{i4kq} \quad (5)$$

where the first summation of g terms is across all relationship effects not involving person 4, and the second summation is across all relationship effects in which person 4 is the target. Note that, unlike the noncomponential estimate, the perceiver and group effects have been removed. Most importantly, the big advantage of the componential measure is the removal of the group effect m_{kq}. Note that the mean is not present in the componential formula but is present in the noncomponential formula.

The key question then is: What do we gain, if anything, by removing the mean? There are two potential advantages of controlling for the group effect. The first advantage is that if there were group variance, removing this variance would increase the reliability of the estimate of the target effect. The second advantage relates back to Malloy et al.'s (1997) primary concern—that the group effect is correlated across contexts. By removing the group effect, we remove this potential confound. These two issues can be addressed using the current data.

The first question is: Does the mean vary across groups? That is, in some families are members seen as intelligent, whereas in other families are members seen as unintelligent? Within the SRM, the question concerns whether family is a source of variance. It can be shown that the theoretical variance in the group means can be estimated as

$$s_{Mq}^2 - (s_{aq}^2 + s_{bq}^2 + 2s_{abq})/4 - (s_{gq}^2 + s_{ggq})/12$$

where s_{Mq}^2 is the variance in the *estimated* group means in context q, s_{aq}^2 is the estimated actor variance for context q, s_{bq}^2 is the estimated partner variance for context q, s_{abq} is the actor-partner covariance for context q, s_{gq}^2 is the relationship variance for context q, and s_{ggq} is the relationship covariance[6] for context q. The estimated group variances for 8 of the 15 group means are negative, which indicates no group variance, and the average of the 15 variances is virtually zero, being only .006. The largest group variance is for judgments of friends on Factor IV (Emotional Stability), but this variance was less than the variance due to target and relationship, and constituted only 10% of the total variance. Thus, there is little indication of any group variance.

Given that there is no group variance, the group means do not estimate a group effect, but rather only reflect chance variance. Therefore, we would expect the group mean to not correlate across contexts. What then are the correlations between the group mean components? The simplest approach is to take the judgments made by all 4 judges within each context, 12 judgments in all, and compute a mean. These means can be correlated across contexts. As seen in Table 3, two of the three cross-context correlations are nonzero. We might then infer that there is a correlation across contexts. However, upon closer examination, we realize that these correlations also contain confounding component effects. The mean within each context contains the perceiver and target effect of person 4. When we compute the mean for each context, person 4 contributes as both a perceiver and a target. To the extent to which person 4 sees others the same way across contexts, or is seen the same way across contexts, the means would then artificially correlate across contexts. To control for this potential confound, we need to recompute the mean eliminating person 4's data. We refer to such means as the *adjusted means*. These cross-context correlations are also presented in Table 3 (again averaged over the Big Five). We see that these correlations are essentially zero. These correlations lead us to the conclusion that results from a noncomponential analysis are not, in fact, confounded by the group effects.

Note that the noncomponential analysis yields an average correlation between target effects of .32, whereas the componential average is .24. Evidently, the greater correlation for the noncomponential approach is not due to the bias of the group means. Why then do the noncomponential and componential correlations differ? By calculating reliability estimates, we

[6]The relationship covariance in interpersonal perception refers to dyadic correlation of perceptions: If i sees j as particularly intelligent, does j see i as particularly intelligent?

can determine that the componential correlations are lower than the noncomponential correlations, namely, because the noncomponential measures are more reliable than the componential ones.

The reliability of the noncomponential measures can be relatively easily obtained. A one-way analysis of variance is computed in which the independent variable is participant, person 4, and the dependent measure is the rating made by each of the three perceivers within a particular context. Therefore the MS_B refers to the mean square between participants and represents agreement of the three judges for the participant on a particular measure. The MS_W represents disagreement between those judges. The estimate of reliability of a *single rater* is the familiar intraclass correlation of

$$\frac{MS_B - MS_W}{MS_B \ + (k-1)MS_W}$$

where k is the number of judges, being 3 in the study under consideration. However, we seek the reliability of the average of three raters and that reliability, what Shrout and Fleiss (1979) referred to as ICC (1, 3), is

$$\frac{MS_B - MS_W}{MS_B}$$

Applying this formula to the Malloy et al. (1997) data, the resulting average reliabilities are .52 for Home, .49 for Leisure, and .61 for Work, the overall average being .54. These reliabilities may look low, but in reality, they are rather typical for interpersonal perceptions. In the most extensive review of peer agreement correlations, Kenny, Albright, Malloy, and Kashy (1994) found that the average correlation between personality judgments for close acquaintances is .275. Given this correlation and using the Spearman-Brown (Nunnally & Bernstein, 1999) prophecy formula, the reliability of three perceivers should be about .53, a value remarkably close to .54.

Given the reliabilities, we can disattenuate the cross-context correlations using noncomponential analysis. Thus, with these reliabilities, we can forecast what the cross-context correlation would be if the measures had perfect reliability. The average cross-context correlation using noncomponential analysis is .32, making the average disattenuated correlation .59 (.32/.54). Thus, persons are seen similarly across contexts, though not exactly the same.

Computing the reliability for the componential measures is not simple. We describe a rather complicated rationale for a formula in the Appendix, which yields an average reliability for the target effect of .49 for Home, .34 for Work, and .48 for Leisure. These reliabilities are lower than the noncomponential reliabilities, which explain in large part why the noncomponential cross-context correlations are larger than the componential. If we use average correlations and reliabilities, the disattenuated cross-context correlation is .24/.44 = .55 for componential analysis and .32/.54 = .59 for noncomponential analysis. Thus, once we adjust for differences in reliability, the componential and the noncomponential correlation give essentially the same result.

What did we learn specifically about the advantages and disadvantages of componential and noncomponential analyses from the Malloy et al. (1997) study? We conclude that people do see a target person in a similar way across contexts, the disattenuated correlation being about .50. There appears to be little or no bias in the noncomponential because there is little or no group variance, and what little variance does exist does not correlate across contexts. Moreover, the componential correlations are lower than the noncomponential correlations, because the noncomponential measures are more reliable than the componential ones.

We agree that, at least in this particular case, we ultimately did not need to bother with the complexity of componential analysis. Malloy et al. (1997) suggested that group means might be correlated across contexts. Although such a suggestion was reasonable, empirical analyses revealed that this concern was only theoretical and not actual. Importantly, there was no way of knowing this beforehand. It was only through executing the componential analysis and addressing potential problems that we learned that componential analysis, at least in this case, produced estimates consistent with noncomponential analysis. The assumption that a component is not present in interpersonal perception is insufficient justification for not conducting a componential analysis.

The presence of a group effect is perhaps not very compelling theoretically. We turn to a third example in which the components have a strong theoretical meaning.

Self-Enhancement

The prior two sections examined correlations between perceptions, and we described what was gained from a componential analysis of those perceptions. In this section, we focus on a concept and show how thinking of the concept in terms of components has theoretical and empirical benefits.

The third example is the question of self-enhancement. Kwan et al. (2004) noted that self-enhancement can be operationalized in two very different ways. The first is based on social comparison theory (Festinger, 1954), which uses the comparison between an individual's self-perception to the perceptions of others to determine the extent to which an individual self-enhances. The second is Gordon Allport's (1937) notion of self-insight, which compares an individual's self-perception to perceptions of that individual by others.

Thus, there are two different ways by which self-enhancement has been conceptualized: Either people can see themselves as better than they see others, or they can see themselves as better than others see them. Self-enhancement would best be examined by combining both theoretical approaches in a componential analysis.

Kwan et al. (2004) proposed that three variables are needed for the measurement of self-enhancement: (a) self-perception, (b) perception of others, and (c) perception by others. If analyzed independently, the social comparison index and the self-insight index each ignore one important component; thus they both portray an inaccurate assessment of self-enhancement bias, and can lead to potential confounds.

Kwan et al. (2004) conceptualized self-perception as a form of interpersonal perception where perceiver and target are the same person. We begin with the SRM model of the perception of one person of another:

$$X_{ijk} = m_k + a_{ik} + b_{jk} + g_{ijk}$$

where m_k is the group mean, a_{ik} is the perceiver effect for person i (how person i generally sees others) in group k, b_{jk} is the target effect (how person j is generally seen by others), and g_{ijk} is the relationship effect from i to j (how person i uniquely sees person j). Following Kwan et al., the model[7] for self-perception of person i of him- or herself is:

$$X_{ii} = o_k + a_{ik} + b_{ik} + h_{iik}$$

where o_k is the mean self-perception in group k, a_{ik} is the perceiver effect for person i (how person i generally sees others), b_{ik} is the target effect (how person i is generally seen by others), and h_{iik} is the "relationship" effect from i to i (how person i uniquely sees him- or herself). It is this last component, h_{iik}, which measures individual differences in self-enhancement. To compute this component, we must subtract from self-perception not only the mean but also the perceiver and target effect. The social comparison approach subtracts only the perceiver effect and the self-insight approach subtracts only the target effect. The SRM approach used by Kwan et al. is to subtract both.

Kwan et al. (2004) empirically tested the major predictions of the model to determine if the social comparison and self-insight indexes yield the same results for self-enhancement. In addition, the authors expanded upon Taylor and Brown (1988) to ask the question: Are individual differences in self-enhancement correlated with individual differences in mental health?

Kwan et al. (2004) studied 128 students who interacted with each other in groups of four or five in a round-robin design and rated each other on 32 personality traits. Three measures of adjustment were also obtained: the Rosenberg Self-Esteem scale (1965), Relationship Harmony (Kwan, Bond, & Singelis, 1997), and Task Performance. When the 32 traits were analyzed, on average 17% of the variance was due to perceiver and 27% of the variance was due to target. Kwan et al. (2004) then averaged perceiver and target effects across the 32 traits to obtain component measures of perceiver and target. They then computed three indexes:

- SRM Index = Self Rating – Perceiver Effect – Target Effect
- Social Comparison Index = Self Rating – Perceiver Effect
- Self Insight Index = Self Rating – Target Effect

Note that all three indexes involve the subtraction of one or more component from self rating. The advantage of the SRM Index is that it combines both approaches by subtracting two components.

When the indexes were correlated with the three adjustment measures, Kwan et al. (2004) concluded:

> When compared with the new SRM index, the social comparison index misleadingly indicated a much stronger self-enhancement effect for self-esteem and no self-enhancement effect (rather than a negative one) for task performance. Similarly, compared with the new SRM index, the self-insight index misleadingly indicated a more positive self-enhancement effect for relationship harmony and less negative self-enhancement effect for task performance. (p. 104)

Thus, more informative results were obtained by removing two components not just one.

The self-enhancement example illustrates the theoretical importance of thinking in terms of components. In measuring self-enhancement, the social comparison approach removed one component and the self-insight removed another. A more comprehensive approach is to remove both.

Criticisms of Componential Analyses

We have seen that there are advantages in measuring components and removing their effects. For the Kenny and Nasby (1980) example, it was only after re-

[7]As discussed in Kenny (1994), the more general model is to allow the actor and partner effects to have weights that do not equal unity.

moving components that we were able to find reciprocity of attraction. For the Malloy et al. (1997) example, we obtained the same results using a componential and a noncomponential analysis. Still, by doing a componential analysis, we learned the measure of consensus was not inflated by group effects. Finally, for the Kwan et al. (2004) example, we reviewed a conceptual refinement of self-enhancement that involved removing two components, the perceiver and the target.

Despite these benefits, several investigators have argued against the necessity of measuring and controlling for components. In this section, we carefully consider the different reasons for not undertaking a componential analysis.

Componential Analysis Is Not Always Possible

It is important to consider that some research designs preclude a componential analysis. Consider for instance, the study of cross-context agreement using a simpler design than our illustrative example in which each participant is judged by just one person in each context. If, for example, we do not have multiple perceivers within the context, we cannot control for the group mean and a componential analysis cannot be undertaken. We should realize that even though a componential analysis is impossible with the simple, single-perceiver design, the components are still present in the judgments, and so the analyses of those judgments may be biased. Malloy et al. (1997) deliberately chose the more complicated design (see Figure 1), because they were concerned about the biasing effects of group mean and perceiver effect components. Therefore, although simple designs preclude a componential analysis, those components are still lurking.

Funder (2001) made the point that researchers need to be careful in implementing study designs in interpersonal perception research. He worried that if, for example, in a cross-context study where we obtain round-robin ratings, the individuals within the social contexts (e.g., the friends) may not know each other well. His point is well taken; obviously designs cannot be blindly applied.

No Interest in Components

Another issue arises when the researcher claims to be interested in only the raw correlation, and computes this correlation without controlling for components. Obviously, if there is interest in only the raw-score correlation, then it makes sense to compute that correlation. However, the meaning of such a correlation is sometimes ambiguous. If a componential analysis is possible, then the correlation can be partitioned into components. Returning to the Malloy et al. (1997) study, it is a very different question to ask whether two people agree in rating the same person (i.e., a correlation of the target components), versus do two people agree when rating different persons (i.e., a correlation of the group mean components). The raw-score correlation contains a mix of both of these correlations. If a researcher wishes only to compute the raw correlation, they need to state the different possible interpretations of such a correlation.

Jussim (2005) claimed that one may simply be interested in whether or not a given perceiver is accurate when making a single prediction:

> But if one wants to determine whether my prediction [that Mike Piazza is the best home-run-hitting catcher] is accurate, the only thing we need to do is figure out whether he hit the ball over the outfield fence in fair territory (in the case of a single judgment), and whether he hit more home runs during his career than any other catcher (in the case of multiple judgments). (p. 63)

Jussim (2005) was indeed correct that in some situations, particularly applied ones, it is important to predict whether a perceiver is right or wrong. For instance, it is a matter of life and death for counselors to be able to determine if a client is suicidal. The science of interpersonal perception has a strong interest in the overall degree of accuracy. Additionally, the science also focuses on the process by which accuracy is obtained by partitioning the overall accuracy into components.

Throwing Out the Baby With the Bath Water

Critics of componential analysis often worry that the "baby is thrown out with the bath water" when componential analyses are done. Consider the previously discussed cross-context study of Malloy et al. (1997). The focus of that study was on how a person was viewed in three different contexts: home, work, and leisure. The noncomponential measure of how a person is seen in one context (e.g., home) would be to simply average how he or she was seen by his or her three family members. The componential measure, see Equation 5, contains that average as well as how the person sees the other family members and the mean of judgments by all four family members of each other. This estimate is a measure of the SRM component of partner (the "baby"), subtracting off the unwanted components of perceiver and group (the "bath water"). However in doing so, error is added which makes the reliability of the componential measure lower than the noncomponential measure. To continue with the metaphor, none of the baby is thrown out with the componential measure, but some distilled water is added after the bath water is thrown out. By removing the bias of the group effect, more error is added into the estimate of the partner effect.

The trade-off between bias and efficiency is well known in the statistical literature. Sometimes, slightly biased estimators are preferred over less efficient ones. Some might wish to argue that a noncomponential analysis is better than a componential analysis because it is more powerful. We think such a position is mistaken. For the Malloy et al. (1997) illustration, there was relatively little bias. If, however, there were even a moderate bias due to including the group mean effect, the noncomponential result would have been both more biased and less efficient. Additionally, very often in componential analyses the variance due to the perceiver (i.e., how the judge generally views others) is removed. Generally, the perceiver accounts for much more variance than the mean and so its removal is much more important. Thus, we think that it is ill-advised to ignore bias just to increase efficiency.

The argument gains credence if a component is not regularly found. For example, many studies of interpersonal perception using intact groups (e.g., families) are conducted and we do not find group effects. Then, based on this body of evidence, a researcher can argue that it is permissible to ignore the group component.

Components Removed by Design Not Analysis

Sometimes components can be removed, not through componential analysis, but by research design. As Funder (2001) discussed, if there are worries about the confounding effect due to perceiver, that effect can be controlled for by asking perceivers to rank order the targets. Many times it is better to choose a design that eliminates a bias rather than removing the bias by statistical analysis.

Which Components?

Jussim (2005) argued that there are difficulties when one considers all of the different componential models: “There is no one right way to divide up components of social perception. … If components are ‘real’ and ‘must’ be assessed, then only complete way to do it would be to assess the more than 50 components” (p. 64). Jussim claimed that given the variety of componential approaches (e.g., Cronbach, 1955; Judd & Park, 1993; Kenny & La Voie, 1984), there are just too many components to ever estimate in one research design.

If it were possible or meaningful to implement all componential approaches in a single study, this concern would be valid. However, one cannot combine SRM and Cronbach partitioning in a single interpersonal perception study simultaneously. The Cronbach method examines the data for each perceiver, whereas the SRM method examines the data for each measure. The Cronbach approach is appropriate if the focus is idiographic, whereas the SRM is appropriate if the focus is nomothetic (Kenny & Albright, 1987; Kenny & Winquist, 2001).

Alternatives to Componential Analysis

Jussim (2005) argued that componential analysis is just one of several approaches to the study of accuracy. He mentioned Brunswik’s (1952) lens model and Funder’s (1995) Realistic Accuracy Model (RAM)[8] as possible alternative models. Earlier we described accuracy in interpersonal perception research as the measurement of the associations between perceptions and a criterion. Both the lens model and RAM start with these two variables, but they in essence specify mediating variables. Thus, these models are not alternatives to a componential analysis; they could be used together.

Consider the interpersonal perception study by Gosling, Ko, Mannarelli, and Morris (2002). The authors had 7 perceivers rate the personality of 83 targets after viewing the targets’ bedrooms. Using Brunswik’s lens model, Gosling et al. specified 42 different possible cues (i.e., mediators[9]) that perceivers may have used. Three perceivers judged the 42 cues for each of the 83 bedrooms, but these perceivers were not the same perceivers as those who made the personality judgments. Each target’s personality was rated by one or two acquaintances.

A componential analysis is possible on the cues that have a two-way structure of Perceiver × Target. However, because there are different perceivers for the judgments, the cues, and the criterion, the target is the only common component. Thus, a mediational analysis can only be performed at the level of the target; this is what in essence Gosling et al. (2002) did. If the same perceivers were used for all the measures, we could perform a mediation analysis at three different levels: the target, the perceiver, and the target × perceiver interaction. Thus, RAM and the lens model are not really alternatives to a componential analysis.

Componential Analyses Are Too Hard To Do

The final criticism is implicit. Some researchers think that componential analysis is very difficult. We agree. Componential analysis can be complex, and clearly, noncomponential analysis is much simpler to

[8]Jussim (2005) also mentioned Dawes’ (1979) improper linear models, which we do not see as model but rather as statistical necessity.

[9]Note that a cue is probably better conceived as a spurious variable; that is, the cue causes both the judgment and the criterion. However, as MacKinnon, Krull, and Lockwood (2000) have shown, spurious variables can be treated as if they were mediators.

do. Consider the measurement of the SRM target effect in Malloy et al. (1997) in Formula 4. In the noncomponential analysis, we simply compute the average of the three informants. In the componential analysis, we take that average, multiply it by 1.125, and then adjust out the average rating made by the target and the grand mean of all judgments made in the group (see Equation 4). Certainly, it is not at all transparent why we multiply by 1.125, nor why we need to adjust out the judgment made by the target of others. However, it is through such a weighting and correction that the componential estimate mathematically removes the perceiver and group effects, components that are present in the noncomponential measure.

The reliability of the noncomponential measure is relatively simple. We compute the intraclass correlation (Shrout & Fleiss, 1979). The derivation of the reliability of the componential measure is very complex (see the Appendix). In fact, we were quite surprised at how involved it was.

A further difficulty is in the measurement of the group variance and covariance. Consider for instance, the test of whether the group effect correlates across contexts. The obvious answer to this question is to correlate the group means. However, as we explained, because the participant serves as a perceiver and a target in all three group contexts, there is a confound. Thus, we needed to compute adjusted means.

As our reanalysis of the Malloy et al. (1997) study shows, componential analysis is indeed very complex. That complexity can sometimes obscure the answer to the original research question. The componential researcher needs to attach conceptual meaning to the mathematics and equations of componential analysis.

Conclusion

What do we conclude generally about componential analysis versus noncomponential analysis? We conclude that despite its complications, it is still advantageous to conduct componential analysis where appropriate. Interpersonal perception data have components of perceiver, target, measure, and their interactions. If the design permits and the research involves correlating sets of perceptions, then the researcher needs to separate the measure into components and compute correlations between the components that make up that measure. When a componential analysis is necessary, the researcher can be confident that the results are not biased by a theoretically irrelevant component.

As we stated earlier, the landscape of interpersonal perception research radically changed after Cronbach (1955). An interpersonal perception score is made up of components. For many questions, researchers should be interested not in the raw score but in the components that make up that measure. Certainly it would be easier if we could simply use raw scores and not go through all of the trouble of componential analysis. However, the complexity of componential analysis, though not always required, can give us a detailed and clear picture of interpersonal perception. All of the trouble is worth it!

References

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.

Brunswik, E. (1952). *The conceptual framework of psychology*. Chicago: University of Chicago Press.

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

Curry, T. J., & Emerson, R. M. (1970). Balance theory: A theory of interpersonal attraction? *Sociometry, 33,* 216–238.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist, 34,* 571–582.

Desoto, C. B., & Kuethe, J. (1959). Subjective probabilities of interpersonal relationships. *Journal of Abnormal and Social Psychology, 59,* 290–294.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117–140.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102,* 652–670.

Funder, D. C. (2001). Three trends in current research on person perception: Positivity, realism, and sophistication. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 319–331). Englewood Cliffs, NJ: Lawrence Erlbaum Associates, Inc.

Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among judges of personality: Interpersonal relations, similarity, and acquaintanceship. *Journal of Personality & Social Psychology, 69,* 656–672.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality & Social Psychology, 82*, 379–398.

Griffin, D., & Gonzalez, R. (1995). Correlational analysis of dyad-level data in the exchangeable case. *Psychological Bulletin, 118,* 430–439.

Israel, J. (1958). A note in measurement of cross-pressures in groups. *Acta Psychologica, 14*, 353–358.

John, O. P. (1990). The "big five" factor taxonomy of personality in the natural language and in questionnaires. In L. A. Pervin, (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York: Guilford.

Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review, 100,* 109–128.

Jussim, L. (2005). *Accuracy in social perception: Criticisms, controversies, criteria, components, and cognitive processes.* In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 37, pp. 1–93). San Diego, CA: Elsevier.

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

Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin, 102*, 390–402.

Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the big five. *Psychological Bulletin, 116*, 245–258.

Kenny, D. A., Bond, C. F., Jr., Mohr, C. D., & Horn, E. M. (1996). Do we know how much people like one another? *Journal of Personality & Social Psychology, 71*, 928–936.

Kenny, D. A., & La Voie, L. (1984). Separating individual and group effects. *Journal of Personality and Social Psychology, 48,* 339–348.

Kenny, D. A., & Nasby, W. (1980). Splitting the reciprocity correlation. *Journal of Personality and Social Psychology, 38,* 249–256.

Kenny, D. A., & Winquist, L. (2001). The measurement of interpersonal sensitivity: Consideration of design, components, and unit of analysis. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 265–302). Englewood Cliffs, NJ: Lawrence Erlbaum Associates, Inc.

Kwan, V. S. Y., Bond, M. H., & Singelis, T. M. (1997). Pancultiral explanations for life satisfaction: Adding relationship harmony to self-esteem. *Journal of Personality & Social Psychology, 73,* 1038–1051.

Kwan, V. S. Y., John, O. P., & Kenny, D. A., Bond, M. H., & Robins, R. W. (2004). Reconceptualizing individual differences in self-enhancement bias: An interpersonal approach. *Psychological Review, 111*, 94–110.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science, 1,* 173–181.

Malloy, T. E., Albright, L., Kenny, D. A., Agatstein, F., & Winquist, L. (1997). Interpersonal perception and metaperception in non-overlapping social groups. *Journal of Personality and Social Psychology, 72,* 390–398.

Newcomb, T. M. (1961). *The acquaintance process*. New York: Holt, Rinehart & Winston.

Norman, W. T. (1967). On estimating psychological relationships: Social desirability and self-report. *Psychological Bulletin, 67,* 273–293.

Nunnally, J. C., & Bernstein, N. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill.

Sherman, S. J., & Wolosin, R. J. (1973). Cognitive biases in a recognition task. *Journal of Personality, 41,* 395–412.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Use in assessing rater reliability. *Psychological Bulletin, 86,* 420–428.

Snyder, M., & Ickes, W. (1985). Personality and social behavior. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., Vol. 2, pp. 883–947). New York: Random House.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103,* 193–210.

Zajonc, R. B., & Burnstein, E. (1965). Structural balance, reciprocity, and positivity as sources of cognitive bias. *Journal of Personality, 33,* 570–583.

Appendix: Derivation of the Reliability of the Partner Effect

As is shown in Formula 5, the estimate of the target effect is:

$$b_{4kq} - \Sigma_{\,i} b_{ikq}/4 - \Sigma\Sigma 0.125 g_{ijkq} + 0.250 \Sigma g_{ijkq}$$

We need to determine the true variance divided by the total variance. The total variance can be determined by computing the variance of the estimated target effects. The true variance is the variance of $b_{4kq} - \Sigma_i b_{ikq}/4$, which can be shown to equal (.5625) σ_{bq}^2 where σ_{bq}^2 is the variance in the target effects for context q. There is one last complication. The target is not randomly sampled from the group and so there might be mean of the target effect may have a nonzero mean. We need to then adjust the estimated target variance by the mean and that adjustment is M_{4q}^2.