

Perceiving Others' Personalities: Examining the Dimensionality, Assumed Similarity to the Self, and Stability of Perceiver Effects

Sanjay Srivastava
University of Oregon

Steve Guglielmo
Brown University

Jennifer S. Beer
University of Texas at Austin

In interpersonal perception, “perceiver effects” are tendencies of perceivers to see other people in a particular way. Two studies of naturalistic interactions examined perceiver effects for personality traits: seeing a typical other as sympathetic or quarrelsome, responsible or careless, and so forth. Several basic questions were addressed. First, are perceiver effects organized as a global evaluative halo, or do perceptions of different traits vary in distinct ways? Second, does assumed similarity (as evidenced by self-perceiver correlations) reflect broad evaluative consistency or trait-specific content? Third, are perceiver effects a manifestation of stable beliefs about the generalized other, or do they form in specific contexts as group-specific stereotypes? Findings indicated that perceiver effects were better described by a differentiated, multidimensional structure with both trait-specific content and a higher order global evaluation factor. Assumed similarity was at least partially attributable to trait-specific content, not just to broad evaluative similarity between self and others. Perceiver effects were correlated with gender and attachment style, but in newly formed groups, they became more stable over time, suggesting that they grew dynamically as group stereotypes. Implications for the interpretation of perceiver effects and for research on personality assessment and psychopathology are discussed.

Keywords: social perception, perceiver effects, informant reports, self-perception, Halo effect

What a most extraordinary child!” Then she frowned. She’s “glad” I punished her, and I “mustn’t feel bad one bit,” and she’s going to “love to live” with me! Well, upon my soul!

—Eleanor H. Porter, *Pollyanna*

In Eleanor Porter’s novel *Pollyanna*, the title character deals with the ups and downs of life by playing the “glad game,” in which she strives to always see the good in every situation and in every person. To a modern psychologist’s eye, the glad game combines several distinct mental processes and attributes, including optimism, reappraisal, and self-efficacy beliefs. The word *Pollyannaish* has evolved in everyday language into a synonym for an extreme and rigidly positive outlook, but psychologists recognize that even within relatively normal populations, people may exhibit some Pollyanna-like tendencies to varying degrees.

In the present investigation, we examined an aspect of Pollyanna’s outlook: tendencies or biases in perceiving other people’s personalities. Specifically, we investigated perceptions of a typical

“other”—tendencies to believe that other people are responsible or careless, supportive or quarrelsome, open-minded or closed. These general perceptions of others—what we call *perceiver effects* (after Kenny, 1994)—are important to understand for a variety of reasons. Methodologically, many studies rely on informant reports as a data source (Kraemer et al., 2003; Vazire, 2006), and a model of perceiver effects can enhance researchers’ understanding of how such reports are formed. Substantively, biases and distortions in perceptions of others have been theorized to be part of a variety of intra- and interpersonal processes, individual differences, and forms of psychopathology (e.g., Benjamin & Wonderlich, 1994; Gara et al., 1993). We focused our investigation on interpersonal perceptions of the personality traits from the Big Five model (John & Srivastava, 1999).

In one cross-sectional study and one longitudinal panel study, we tested a series of interrelated questions about perceiver effects. First, are perceiver effects organized as a global evaluative halo, or are perceptions of different traits affected by different factors? Second, is assumed similarity (believing that others have similar traits as the self) trait specific, or does it simply reflect a broad similarity-of-evaluative tone? Third, do perceiver effects reflect a stable concept of the generalized other, and/or are they reshaped in relation to new social contexts as group-specific stereotypes? We begin by more formally defining the perceiver effect, and then we review theories and evidence related to our three central questions.

Perceiver Effects: Definitional and Conceptual Issues

Our definition of perceiver effects comes from the social relations model (SRM; Kenny, 1994; Kenny & La Voie, 1984). The

Sanjay Srivastava, Department of Psychology, University of Oregon; Steve Guglielmo, Department of Psychology Brown University; Jennifer S. Beer, Department of Psychology, University of Texas at Austin.

Data collection for Study 2 was funded in part by a Sigma Xi grant in aid of research awarded to Sanjay Srivastava and by a University of California, Berkeley, Sigma Xi chapter grant and a Graduate Women in Science grant awarded to Jennifer S. Beer.

Correspondence concerning this article should be addressed to Sanjay Srivastava, Department of Psychology, 1227 University of Oregon, Eugene, OR 97403-1227. E-mail: sanjay@uoregon.edu

SRM is a conceptual and empirical model of the factors that contribute to interpersonal perceptions. Suppose that we ask Iris how friendly she thinks Jack is, and Iris replies that Jack is moderately friendly—say, a 7 on a scale ranging from 1 to 10, where the average rating is a 5. Within the SRM, there are four possible explanations (not mutually exclusive) why Iris perceives Jack as moderately friendly. The average rating of 5 is the starting point that tells us how commonly friendliness is perceived, regardless of anything particular about Iris or Jack. Beyond that, Iris might be unusually likely to see other people as friendly, perhaps because of her assumptions about human nature. Jack might be unusually likely to be seen by others as friendly, perhaps because he typically behaves in a friendly way. Or Iris might have a unique perspective that she applies in perceiving Jack, perhaps because he reminds Iris of her brother Mike, who is very friendly. Formally, the SRM may be expressed in the following equation:

$$Y_{i,j} = G + P_i + T_j + R_{i,j}. \quad (1)$$

In this equation, $Y_{i,j}$ is the perception that person i has of person j . In this example, it would be 7, Iris's perception of Jack's friendliness. It is the sum of four components. G is the average perception in the group of perceivers and targets under consideration, which would be 5 in this example. P_i is the perceiver effect associated with perceiver i . In this example, if Iris is unusually likely to see others as friendly, then P_{Iris} would be positive. T_j is the target effect associated with target j ; if Jack is unusually likely to be seen by others as friendly, T_{Jack} would be positive. $R_{i,j}$ is the relationship effect uniquely associated with perceiver i perceiving target j . If Iris has a special or unique perspective on Jack that makes her think he is unusually friendly, $R_{Iris,Jack}$ would be positive. ($R_{i,j}$ also includes measurement error in this example; separating true relationship effects from error requires the use of multiple indicators.)

It is important to emphasize that the perceiver effect is defined in the SRM as a perceptual tendency applied to everyone else in a group: If it is a perceiver effect that leads Iris to say that Jack is very friendly, then she will also say that Ken and Laura are very friendly. Different perceptions of different others, such as if Jack (but not Ken or Laura) reminds Iris of her friendly brother Mike (cf. Chen & Andersen, 1999), are relationship effects in the SRM and are not the focus of the present investigation. We return to this distinction in discussing halo effects and assumed similarity.

In practice, the values of these components cannot be known for a single interpersonal perception considered in isolation: If all we know is that Iris thinks Jack is friendly, we cannot determine how much this perception is due to a high base rate, something about Iris as a perceiver, something about Jack as a target, or the unique perspective Iris has in perceiving Jack. To estimate the four variables, researchers typically use designs in which multiple perceivers rate multiple targets. A special case is the round-robin design for small groups, in which each group member rates each other member; thus, each person is simultaneously a perceiver of others and a target of others' perceptions (Warner, Kenny, & Stoto, 1979).

Examining the Structure of Perceiver Effects

In studies of interpersonal perception, perceiver effects typically account for a substantial portion of the variance in trait ratings

(Kenny, 1994). However, perceiver effects have received relatively less attention than other components of interpersonal perception. Different theories of social perception lead to different predictions about the dimensionality of perceiver effects; thus, one goal of this research was to compare different theory-based structural models.

Perceiver Effects as Differences in Acquiescent Responding

One possibility is that perceiver effects could result from a methodological artifact. Researchers have long been concerned about whether subjects' responses on rating scales are affected by response sets (e.g., Block, 1965). Perceiver effects could be caused by variance in *acquiescence*, defined as "the tendency of an individual to agree or say 'yes' to personality inventory statements, regardless of the content of the item" (Block, 1965, p. 1). According to this explanation, Iris may endorse the item "This person is friendly" not because of some meaningful tendency to perceive others as friendly, but simply because she tends to agree with inventory items. Had Iris instead been asked to rate the statement, "This person is unfriendly," she would have agreed with that statement and seemingly exhibited the opposite perceptual tendency.

The classic psychometric approach to eliminating acquiescence from measures is to construct multiitem scales with a balanced key (items representing both the high and low ends of a continuum). When the SRM has been applied to such scales, a substantial perceiver effect remains, indicating that the perceiver effect cannot be explained away as mere acquiescence (Kenny, 1994). However, acquiescence effects may be present to some modest degree in individual perceptions. Therefore, where possible we attempted to model acquiescence effects and control for them when evaluating substantive models of perceiver effects.

Three Theoretical Models of Perceiver Effects

If perceiver effects are not principally caused by acquiescence, how best can they be explained? One way to address this question is to examine the structure of perceiver effects for different traits—to ask whether or how different perceiver effects go together. If Iris generally sees others as extraverted, should we expect that Iris also sees others as agreeable and conscientious—or are perceiver effects for different traits shaped by different factors? We review three different theoretical models of the perceiver effect: a one-factor global evaluation model, a two-factor model based on core interpersonal themes, and a five-factor model based on the Big Five model of personality traits.

Global Evaluation

One possibility is that perceiver effects are driven by a global evaluation factor—viewing others along a continuum from "good" to "bad." This model would reflect a kind of *halo effect*, defined as "the influence of a global evaluation on evaluations of individual attributes of a person" (Nisbett & Wilson, 1977, p. 250). Research on halo effects has indicated that once a perceiver makes a broad judgment that another person is good (or not), perceptions of specific attributes will be formed in an evaluatively congruent

direction—even in the face of relevant evidence to the contrary (Nisbett & Wilson, 1977). All of the Big Five trait dimensions have a more-desirable pole and a less-desirable pole (John & Robins, 1993; Saucier, Ostendorf, & Peabody, 2001), and therefore could be susceptible to halo effects.

Evaluative halos can affect different components of the SRM. If Iris decides that Jack is good and Ken is bad, and therefore attributes all positive traits to Jack and all negative traits to Ken, then that would indicate a halo in the relationship effect. If Iris decides that all others are good and Ken decides that all others are bad, and therefore Iris attributes all positive traits to all others and Ken attributes all negative traits to all others, then that would indicate a halo in the perceiver effect. It is the latter that is the focus of the present investigation. If perceiver effects for different traits are all driven by a global evaluative halo, then the perceiver effect should be well described by a one-factor structure.

Agency and Communion

Two themes that recur in many theories of interpersonal life are acting energetically to assert or expand the self and establishing connection and union with others. Noting the recurrence of these broad themes across the social sciences, Bakan (1966) proposed the labels of *agency* and *communion*. These two interpersonal themes appear in many areas of psychology, including power and intimacy motivations (McAdams, 1985), status and love concerns in relationships (Foa & Foa, 1974), dominant and affiliative interpersonal orientations (Wiggins, 1991), competence and warmth in stereotype content (Fiske, Cuddy, Glick, & Xu, 2002), and strategies for “getting ahead” and “getting along” in the social world (Hogan, 1983).

Several recent studies indicate that these two themes may also describe a higher order structure of personality or personality perceptions. Although the Big Five trait factors have often been described as orthogonal dimensions, empirical studies of self- and other-reports typically show modest but replicable correlations between some of the factors. Digman (1997) proposed that these correlations yield a superordinate two-factor structure reflecting the two core themes. Subsequently, the “Bigger Two” structure has been replicated by a number of investigators (DeYoung, 2006; Markon, Krueger, & Watson, 2005; Saucier & Goldberg, 2001). In these models, an agency factor represents common features of Extraversion and Openness: acting energetically upon the world and expanding one’s mental horizons. A communion factor represents common features of Agreeableness, Conscientiousness, and low Neuroticism: caring for others, fulfilling responsibilities, and keeping inappropriate emotional impulses under control.

Concerns with agency and communion may potentially drive the assumptions and beliefs that perceivers hold or form about others. In a multi-informant, multitrait analysis of interpersonal perceptions, Biesanz and West (2004) found evidence that the Bigger Two correlational structure is driven primarily by perceivers’ biases rather than by the true co-occurrence of traits. Another study found that a similar two-factor model fit biases in self-perception better than a one-factor model (Paulhus & John, 1998). On this basis, we considered the Bigger Two to be a plausible model for the structure of perceiver effects.

The Big Five Factors

A third theoretical proposition is that the structure of perceiver effects may parallel the structure of personality traits. Research in English-speaking North American samples (and several other linguistic and national groups) has shown that personality trait structure is organized around five broad factors, known as the Big Five (Goldberg, 1990; John & Srivastava, 1999) or the five-factor model (Digman, 1990). This structure replicates across many different methods, and substantial evidence has supported the validity of self- and other-reports of the five factors (Ozer & Benet-Martínez, 2006; Vazire, 2006). However, perceptions of others on all of the Big Five traits contain additional variance beyond an accurate kernel of truth (Kenny, 1994, 2004). In order for people to form accurate perceptions of the five factors, different inferential processes are necessary to meaningfully distinguish among them. These inferential processes may each be vulnerable to different influences. If so, then we might find a five-factor structure in perceiver effects similar to that found in raw self- and other-reports.

Perceiver Effects and the Self: Trait Specificity in Assumed Similarity

Psychologists have long been interested in how perceptions of the self relate to perceptions of others (Cronbach, 1955; Festinger, 1954). At a descriptive level, *assumed similarity* refers to the belief that other people are similar to the self. General assumed similarity would be evidenced by a positive correlation between self-perceptions and perceiver effects. (A belief that the self is similar to some and dissimilar to others, as in Ames, 2004, would affect correlations with relationship effects.) Several studies in which the SRM has been used have found evidence of general assumed similarity—that is, correlations between self-perceptions and perceiver effects—in personality trait perceptions. The correlations are typically positive even at minimal acquaintance, are evident for all of the Big Five, and tend to increase in magnitude as people get to know one another better (Kenny, 1994; Park & Judd, 1989).

If there is assumed similarity, then where does it come from? Kenny (1994) interprets the self-perceiver correlation as originating from enhancement processes that are applied to perceptions of both self and others (see also Kwan, John, Kenny, Bond, & Robins, 2004). Previous studies have not tested the structure of these enhancement processes, but Kwan et al. (2004) combined multiple traits by keying them in the same evaluative direction and averaging them together—an approach that implicitly assumes that enhancement processes operate on global evaluative variance.

Building on the structural theories presented in the previous section, we asked whether assumed similarity operates at the level of broad evaluations or is trait specific. Broad evaluative similarity would mean that people assume that the self and others are similar in evaluative tone (I am good and so are others), whereas trait-specific similarity would be more focused (if I am extraverted, then so are others, regardless of how agreeable I am or they are).

Origins and Formation of Perceiver Effects

Do perceiver effects reflect stable and global beliefs about human nature, or are they reformulated in relation to each new

social context? Impressions of other people are typically organized into enduring memory structures (Smith, 1998). Theories of mental representations in relationships have suggested that individuals have a general schema of what others are like, as well as schemas about groups and particular interaction partners (Baldwin, Keelan, Fehr, Enns, & Koh-Rangarajoo, 1996; Smith, 1998). Thus, as a starting point, we expected to observe some degree of stability within a specific context.

Kenny (1994) proposed two kinds of mental representations that could give rise to stable perceiver effects: a *generalized other*, and *group-specific stereotypes*. The former theory originates with Mead (1934), who proposed that through early social experiences, individuals formulate a concept of the "generalized other" containing beliefs about how other people typically behave. Similarly, attachment research proposes that individuals have stable working models of relationships that may be applied when interacting with others (Fraley, 2002). According to both accounts, mental representations of the generalized other are stable by adulthood. If perceiver effects are based on preexisting beliefs that individuals bring to new relationships, then they should be related to stable, preexisting individual differences. We focused on two individual differences, gender and attachment style. Gender has been associated with perceiver effects in previous studies, such that women's perceiver effects of various traits are in a more evaluatively positive direction (Winqvist, Mohr, & Kenny, 1998). Attachment theory proposes that attachment styles are rooted in stable working models of relationships (Bartholomew & Horowitz, 1991; Fraley, 2002). We did not have a basis for making strong trait-specific predictions, but we expected that secure individuals would have more evaluatively positive perceiver effects and insecure individuals more negative perceiver effects.

Another possibility is that perceiver effects may stem from group stereotypes (Kenny, 1994). According to this view, as individuals join and get to know new groups, they form local expectations about what that particular group is like. Because group-specific stereotypes are formed anew for each new group, the perceiver effect would become increasingly stable over time as perceivers form a clearer mental representation of their group.

The generalized-other theory and the group-stereotypes theory are not mutually exclusive, but they lead to different (noncompeting) hypotheses about the stability of perceiver effects over time in intact groups. The generalized-other theory makes predictions about preexisting individual differences and about the initial level of stability in a new group, whereas the group-stereotypes theory makes predictions about how stability increases over time. If both effects are in operation simultaneously, perceiver effects will be moderately stable early in the life of a group, and then become even more stable over time.

Overview of the Present Studies

In two studies, we examined the structure of perceiver effects, their relations to the self, and their stability over time. All of the studies involved naturalistic interactions between subjects who interacted in small groups. Nearly all subjects began the study as strangers, and each subject's perceiver effects were derived relative to other group members. This allowed us to ensure that interpersonal perceptions were not based on social interactions outside of the experimental setting: We could rule out preexisting

differences in objective qualities of the relationships by studying subjects who were initially strangers.

In Study 1, a large sample of subjects played an icebreaker game and then reported trait perceptions. We used this study to examine the structure of perceiver effects by comparing the relative fit of nested models specifying acquiescence, global evaluation, agency and communion, or the Big Five. We also examined correlations between self-perceptions and perceiver effects to see whether we could replicate evidence of assumed similarity, and we controlled for self-perceptions and perceiver effects of other traits to test trait specificity. Study 2 was a longitudinal panel study in which subjects came to the laboratory four times to interact with the same group of people. We used this study to replicate the model comparisons from Study 1. We again tested for self-perceiver correlations and controlled for perceptions of other traits. To look at the origins and formation of perceiver effects, we examined correlations with stable individual differences, and we took advantage of the longitudinal design to examine the stability of perceiver effects over time to see whether stability increased as subjects became better acquainted.

Study 1

One of the main goals of Study 1 was to compare different models of the perceiver effect. Subjects were strangers who played a brief 5-min icebreaker game together. Previous studies have shown that subjects in such low-acquaintance designs can form modestly accurate perceptions, but a sizable perceiver effect remains (Albright, Kenny, & Malloy, 1988). Subjects rated each other's personalities on 10 items representing the high and low poles of each of the Big Five. Using confirmatory factor analyses (CFAs), we began the model comparisons with a baseline model that specified only an acquiescence factor that was fixed to load positively on all items (i.e., in the direction of greater agreement, regardless of item content). We then tested a series of nested models that specified an acquiescence factor plus one, two, or five substantive factors corresponding to global evaluation, agency and communion, or the Big Five.

Method

Subjects and procedure. The 423 subjects in this study were students in introductory psychology classes who took part in a classroom demonstration about first impressions. The procedure was conducted in two different classes, and data were combined for the present study. Demographic data were not collected; however, from other studies done in similar classes, it was estimated that the sample was 65%–70% female, 78%–80% White, 7%–10% Asian and Pacific Islander, 2%–5% Hispanic, and about 8%–10% various other ethnicities.

The procedure was conducted at the first meeting of the class discussion section. Subjects were instructed to form groups of four with people they did not already know (when class sizes were not divisible by 4, groups of three or five were allowed so that everyone could be in a group). Every individual created a fake "code name," and each group made up a group name; these names were written on the questionnaires so that self- and other-reports could be reported and matched anonymously. The subjects then played the game "Two Truths and a Lie." In this game, every

person makes three statements about him- or herself, two of which are true and one of which is not. Other members of the group try to guess which statement is the “lie.” After every group member made their three statements, but before anyone revealed which statement was the lie, each subject rated all other group members (i.e., made round-robin peer ratings) and made self-ratings of personality traits.¹

Measures. Perceptions of others and self-perceptions of Big Five personality traits were measured with the Ten-Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003). The TIPI measures each of the Big Five personality factors with a pair of items, one of which is reverse coded. For example, Agreeableness is measured with the items “Sympathetic, warm” and “Critical, quarrelsome” (reverse coded). Ratings were made on a 7-point scale ranging from 1 (*disagree strongly*) to 7 (*agree strongly*). Because each of the Big Five factors covers a broad content domain (John & Srivastava, 1999), the items for these “scales” were selected to maximize content validity. This means that by design, interitem correlations and internal-consistency tests of reliability, such as Cronbach’s alpha, are expected to be quite modest. However, the TIPI scales have shown substantial test-retest reliability when used as self-report measures of personality (Gosling et al., 2003). Item intercorrelations are reported in the Appendix.

Data reduction. Perceiver effects were calculated from the round-robin ratings using a formula derived by Warner et al. (1979; see also Kenny, 1994). In a round-robin design, a simple average of person i ’s ratings of multiple others is biased, because each member of a group rates a slightly different set of other people (i.e., in a group with Iris, Jack, and Ken, Iris’s perceiver effect is based on her ratings of Jack and Ken; Jack’s perceiver effect is based on his ratings of Iris and Ken; etc.). Therefore, the formula includes an adjustment for the fact that the person making the rating—that is, the perceiver—is missing from the set of individuals being rated, which allows for direct comparisons of different individuals:

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

In this formula, n is the group size, M_i is person i ’s mean rating of all other group members, $M_{.i}$ is the mean rating of i made by other members of the group, and $M_{..}$ is the mean rating in the group as a whole. A perceiver effect was computed for each subject on each item, reflecting the subject’s tendency to endorse that item when rating others. Each subject’s perceiver effect was calculated relative to his or her fellow group members (group-mean centered). Self-reports were also group-mean centered. This made the perceiver effects and self-reports appropriate for analyses that assume independent data. Variances and reliabilities of the perceiver effects are shown in Table 1; reliabilities were calculated using the formula derived by Kenny, Lord, and Garg (1986).²

Analyses. Four distinct and increasingly complex CFA models of the perceiver effects were tested using Mplus (Muthén & Muthén, 2007). The first model specified only acquiescence effects. It was not expected that this model would fit particularly well, but it served as a baseline for specifying and evaluating models with substantive factors. Further models retained the acquiescence factor and added one, two, or five factors correspond-

ing to the a priori theoretical models. The acquiescence factor was always specified as orthogonal to the substantive factors, but substantive factors in the two- and five-factor models were free to covary with one another.

Our model-testing strategy emphasized comparisons of the relative fit of these competing a priori models. One useful basis for comparing relative fit of nested models is chi-square difference tests. Chi-square tests are well understood, and they have the advantage of allowing hypothesis testing with familiar criteria (p values). However, chi-square values are almost always at least marginally smaller (i.e., better) in more complex models, and with adequate power even small differences in fit are often significant. Therefore, fit statistics that are not sensitive to sample size and that penalize for the number of free parameters were also examined, balancing fit against parsimony. The comparative fit index (CFI), root-mean-square error of approximation (RMSEA), and the Bayesian information criterion (BIC) all include penalties for the number of free parameters; BIC in particular tends to strongly reward parsimony. CFI ranges from 0 to 1, and higher values indicate a preferred model. RMSEA has a minimum of 0, and lower values indicate a preferred model. BIC has no standardized scale, so it is not a meaningful index of absolute fit; for relative comparisons of models of the same data, lower values of BIC indicate a preferred model. The standardized root-mean-square residual (SRMR) is also reported, which ranges from 0 to 1, with lower values indicating better fit (though we note that like chi-square, SRMR does not adjust for parsimony). Because the SRM calculations removed group means from the data prior to the CFA analyses, all fit statistics were calculated with an effective sample size $(N - g + 1) = 311$, where g = the number of groups.

Several considerations led us to place less emphasis on absolute fit in making judgments about models. Model fitting needs to balance statistical criteria like goodness of fit against scientific considerations like interpretability and parsimony (Browne & Cudeck, 1993). In the present context, it was unrealistic to expect very good absolute fit. CFA models of self-reported personality traits are robustly replicable and have good relative fit in comparison to alternative specifications, but they only show modest fit in absolute terms (McCrae, Zonderman, Costa, & Bond, 1996). This is because the simple structure specified in most CFA models does not allow cross-loadings and correlated residuals. The five-factor structure of self-reported traits is widely accepted by broad scientific criteria (John & Srivastava, 1999), and it would be unrealistic to expect perceiver effects to fit a CFA model better than self-reports do. Therefore, to provide context for interpreting the ab-

¹ The structure of the game and the timing of when subjects revealed which statement was the playful lie (i.e., after others made ratings) was part of the pedagogical purpose of the demonstration, which was to discuss counterintuitive aspects of impression formation. Substantial previous research (e.g., Ambady & Rosenthal, 1992; Gosling, 2008) has shown that perceivers implicitly base their impressions on a wide variety of cues, such as appearance and nonverbal behavior, that would be unaffected by the truthfulness of the verbal content in this game. Indeed, in this data set “accuracy” correlations (self-target agreement) were in line with previous research on minimal-acquaintance interactions.

² Statistics for the other SRM components (target and relationship variance and reliability) for both studies are available from Sanjay Srivastava upon request.

Table 1
Variances (and Reliabilities) of Perceiver Effects

Trait	Study 1		Study 2			
	Item 1	Item 2 (reversed)	Week 1	Week 2	Week 3	Week 4
Extraversion variance (Reliability)	0.23 (.36)	0.09 ^a (.11)	0.79 (.69)	0.53 (.61)	0.67 (.66)	0.92 (.78)
Agreeableness variance (Reliability)	0.29 (.45)	1.10 (.74)	0.68 (.71)	0.78 (.78)	0.78 (.73)	1.01 (.80)
Conscientiousness variance (Reliability)	0.33 (.48)	0.41 (.46)	0.79 (.75)	0.69 (.73)	1.09 (.82)	1.09 (.86)
Neuroticism variance (Reliability)	0.58 (.56)	0.41 (.50)	2.13 (.85)	2.18 (.83)	2.33 (.84)	2.65 (.88)
Openness variance (Reliability)	0.21 (.35)	0.37 (.42)	0.38 (.61)	0.59 (.67)	0.75 (.70)	1.25 (.83)

Note. $N = 423$ for Study 1 and $N = 152$ for Study 2. All variances are significantly different from zero at $p < .05$.

^a Variance is not significantly different from zero.

solute fit of the models of perceiver effects, the same models were fit to self-perception data so that standardized fit statistics could be compared.

Results and Discussion

Global evaluation, agency and communion, or five factors?

Acquiescence only. Acquiescence bias is defined as a tendency to agree (vs. disagree) with items regardless of their content. Therefore, we specified an acquiescence factor by fixing loadings on all items to +1.0, regardless of their content or keying. As expected, this model showed poor fit (see Table 2). However, the acquiescence factor had significant variance ($\tau = 3.11$, $p < .01$), suggesting that we would be justified in retaining an acquiescence factor in the substantive models.

Global evaluation. The second model added a global evaluation factor by specifying a single factor that was free to load on all items, over and above the fixed acquiescence factor. This model's fit was much better than the acquiescence-only model. The chi-square difference test showed a significant improvement in fit, $\Delta\chi^2(10) = 403.6$, $p < .001$. As shown in Table 2, all of the fit statistics were dramatically better, even with penalties for the additional free parameters in this model. The signs of all 10 item loadings were consistent with the interpretation of a global evaluation factor.

Agency and communion. The next model we considered was the agency/communion model. In this model, we specified an agency factor that loaded on extraversion and openness items, and a communion factor that loaded on Agreeableness, Conscientious-

ness, and Neuroticism items (cf. Wiggins, 1991). The model allowed no item cross-loadings. We also retained the acquiescence factor with fixed loadings. This model fit better than the global evaluation model, $\Delta\chi^2(1) = 26.3$, $p < .001$. As shown in Table 2, all of the fit statistics indicated an improvement, even adjusting for parsimony. The signs of the item loadings were all in the theoretically expected direction.

Five factors. The fourth model we considered included separate factors for each of the Big Five. Items were free to load on their a priori factors but with no cross-loadings. This model also retained the acquiescence factor. This model showed a significant improvement in fit over the agency/communion model, $\Delta\chi^2(9) = 43.5$, $p < .001$. All of the parsimony-adjusted fit indices indicated that this model was preferred over the two-factor model in spite of its greater complexity, although the improvements were less dramatic this time. All loadings were in the expected direction. Absolute standardized loadings ranged from .42 to .86, and all were significant at $p < .001$.

Some of the factors were quite highly correlated (absolute factor intercorrelations ranged from .25 to .93; see Table 3 above the diagonal). Therefore, although a simple global evaluation model did not fit the data, we tested whether a higher order global evaluation factor could explain the intercorrelations among the latent factors. The second-order factor model, depicted in Figure 1, was a reasonable fit; although the chi-square was significantly worse than the five-factor model with free correlations between all factors, CFI and SRMR became a little worse, and BIC and RMSEA showed modest improvements favoring the model with

Table 2
Confirmatory Factor Models of Perceiver Effects in Study 1

Model	χ^2	df	CFI	BIC	RMSEA	SRMR
1. Acquiescence only	566.0 ^a	44	.02	7347	.195	.213
2. Global evaluation plus acquiescence	162.4 ^b	34	.76	7001	.110	.083
3. Agency and communion plus acquiescence	136.1 ^c	33	.81	6980	.100	.079
4. Five factors plus acquiescence	92.6 ^d	24	.87	6989	.096	.063
5. Five factors plus acquiescence; trait factors load on a higher order global evaluation factor	109.5 ^e	29	.85	6977	.094	.069

Note. $N = 423$. Because the variables were group mean-centered, statistics were calculated with the sample size as $(N - g + 1)$, where g = the number of groups. Chi-square values with different superscripts are significantly different from each other at $p < .001$. CFI = comparative fit index; BIC = Bayesian information criterion; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual.

Table 3
Latent Factor Correlations From the Five-Factor Models in Studies 1 and 2

Factor	E	A	C	N	O
Extraversion (E)	—	.37	.25	-.39	.74
Agreeableness (A)	.69	—	.88	-.88	.93
Conscientiousness (C)	.70	.73	—	-.63	.74
Neuroticism (N)	-.31	-.41	-.49	—	-.77
Openness (O)	.72	.73	.81	-.37	—

Note. Correlations between latent trait factors in Study 1 are above the diagonal; those for Study 2 are below the diagonal. All correlations are significant at $p < .05$.

the higher order factor.³ The loadings of the latent trait factors on the higher order evaluation factor ranged from .47 (Extraversion) to 1.0 (Agreeableness). Even with the global evaluation factor, the specific factor variances for Extraversion, Conscientiousness, and Neuroticism were significantly different from zero. Thus, we concluded that perceiver effects reflect a global evaluation factor plus trait-specific factors for at least some traits. The global evaluation factor by itself was not a complete explanation.

Investigating model fit. The strategy of comparing relative fit of a priori models led us to conclude that a five-factor model with a higher order evaluation factor fit the data well. However, in absolute terms, the fit of this model was borderline; for example, CFI was lower and RMSEA was higher than conventional cutoffs for “good” fit. We therefore took two additional steps to help us understand this model. First, we ran identical models on self-report data. Because a five-factor model is widely accepted as the best model of trait self-reports, this would give us a basis for comparing absolute fit statistics. Second, we investigated the modification indices and made post hoc modifications to see whether we could improve model fit.

To provide a context for interpreting absolute fit statistics, we ran the same series of four nested models on the self-report data. As expected, the five-factor model was preferred by all criteria. In comparing standardized fit statistics, the absolute fit of the model of self-reports was quite similar to the model of perceiver effects. For the five-factor model of self-reports, CFI = .85 (compared with .87 in the model of perceiver effects), RMSEA = .096 (compared with .096), and SRMR = .074 (compared with .063). Thus, an a priori five-factor model fit the perceiver effects about as well as it fit the self-reports.

Inspection of the modification indices for the five-factor model of perceiver effects showed that allowing some cross-loadings would improve fit. The largest modification index indicated that we could improve fit by allowing one of the neuroticism items to load on the Extraversion factor. After we made this modification, the indices indicated that we could make a further improvement by allowing an openness item to load on Extraversion. The resulting model showed improved fit over the five-factor model, $\Delta\chi^2(2) = 44.5, p < .01$, CFI = .95, BIC = 6,955, RMSEA = .062, RMR = .04. The new item loadings were -.55 and .40. Although the modifications were post hoc and therefore model fit should be viewed with some caution (MacCallum, Roznowski, & Necowitz, 1992), the modifications strongly suggested that small changes to the model could bring it up to conventional criteria of good absolute fit.

Assumed similarity. Was there evidence of assumed similarity? We formed two-item composites for each of the Big Five traits, and we computed correlations between perceiver effects and self-reports. As shown in Table 4, all of the correlations were positive and statistically significant, indicating that the more a person described him- or herself as having a trait, the more likely the person was to see the trait in others. Across traits, the average assumed similarity correlation was .28.

To test whether there was evidence of trait-specific assumed similarity, we computed partial correlations; for each partial correlation, we controlled for self-perceptions and perceiver effects of all other traits. These correlations should control for any higher order evaluative variance or other factors that could be shared with other traits. The partial correlations, shown in Table 4, were significant for every trait except Conscientiousness; the average partial correlation, .20, was somewhat lower than the average zero-order correlation.

Summary. In summary, model comparisons showed strong support for a five-factor structure over a single global evaluation factor or two agency and communion factors. This was the case even when taking the greater complexity of the five-factor model into account. Factor loadings in the global evaluation and agency/communion models were theoretically consistent, and each was preferred over a simpler, nested model, suggesting that these models have some merit; and a model with five trait factors plus a higher order global evaluation factor fit well. However, the simpler models did not completely capture the pattern of effects, even when weighing the competing concerns of fit versus parsimony; and the higher order factor did not explain all of the variance in the trait-specific factors. From this evidence, we initially concluded that perceiver effects are multidimensional: Perceiver effects for different traits are distinguishable from each other, a finding that we attempted to replicate in Study 2. With regard to assumed similarity, all five traits showed positive self-perceiver effects. Most of the partial correlations, which should have removed any effects of global evaluation, remained significant, suggesting that assumed similarity reflects more than just broad evaluative tone.

Study 2

For Study 2, we expanded on Study 1 in several ways. Study 2 was a longitudinal study: Subjects were assigned to small groups at the start of the study, and each group met and interacted four times. Self-perceptions and perceptions of others (used to calculate perceiver effects) were assessed at the end of each interaction.

As in Study 1, we tested nested models of the structure of perceiver effects. However, this time we approached the question in a different way, asking: What is the stable structure of perceiver effects? In other words, are biases in perceptions of others stable over time; and if so, what is the structure of the stable biases? To address this question, we fit models to the longitudinal data, using perceiver effects from different time points as indicators of the latent factors.

We again examined assumed similarity, testing both zero-order correlations between self-perceptions and perceiver effects, and we

³ A model with two higher order factors, for agency and communion, did not converge (though we note that in Study 2, the model with two higher order factors showed worse fit).

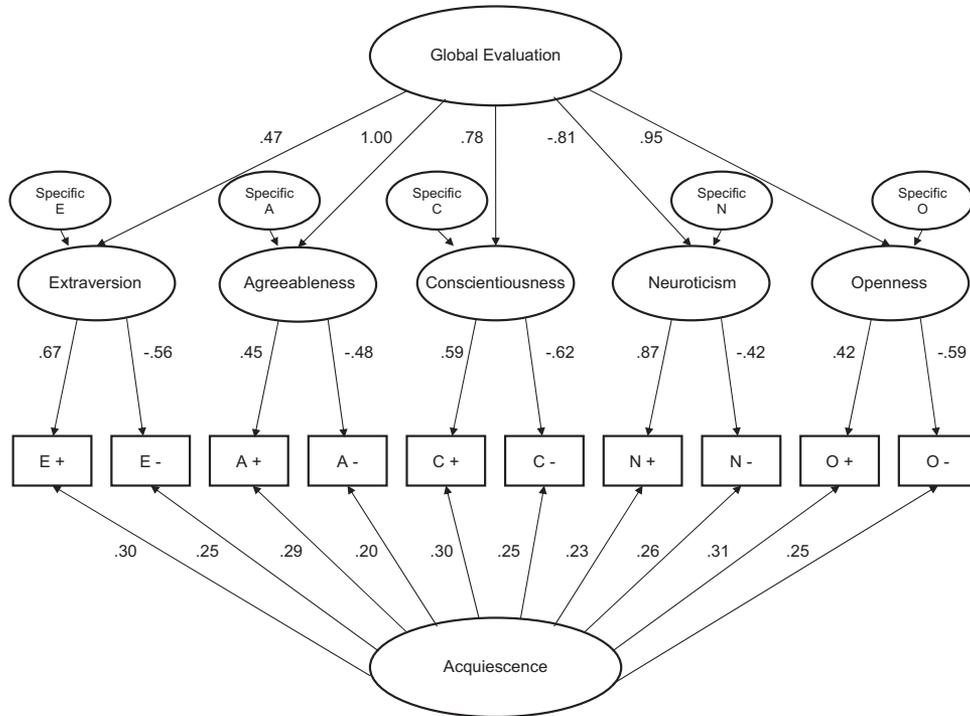


Figure 1. Path diagram depicting the five-factor plus acquiescence model with a second-order global evaluation factor, Study 1. Numbers next to arrows are loadings from the standardized solution. Residuals for the measured variables are not shown.

computed partial correlations for each trait that controlled for self-perceptions and perceiver effects of other traits in order to remove evaluative variance that may be shared between different traits.

Table 4
Assumed Similarity: Correlations and Partial Correlations Between Self-Perceptions and Perceiver Effects

Trait	Study 1	Study 2			
		Week 1	Week 2	Week 3	Week 4
Extraversion <i>r</i>	.20	.31	.22	.45	.44
<i>pr</i>	.14	.29	.15	.26	.42
Agreeableness <i>r</i>	.45	.05	.16	.25	.40
<i>pr</i>	.33	.01	.02	.03	.20
Conscientiousness <i>r</i>	.18	.20	.09	.26	.37
<i>pr</i>	.09	.17	-.04	.03	.20
Neuroticism <i>r</i>	.29	.12	.24	.16	.08
<i>pr</i>	.18	.10	.19	.10	.13
Openness <i>r</i>	.25	.07	.16	.26	.30
<i>pr</i>	.15	.07	.14	.17	.20
Mean <i>r</i>	.28	.15	.17	.28	.32
Mean <i>pr</i>	.20	.11	.09	.11	.23

Note. *N* = 423 for Study 1, and *N* = 152 for Study 2. Numbers in italics are significantly different from zero at *p* < .05. For all partial *r*s, self-perceptions and perceiver effects of all other traits were partialled out. Means were computed using Fisher's *r*-to-*z*' transformation and appear in boldface type.

We tested hypotheses about the origins and formation of perceiver effects in two ways. First, we examined correlations with gender and attachment style to see whether these stable individual differences could predict the perceiver effect as predicted by the generalized-other theory. Second, we examined the stability of the perceiver effects. We specifically tested for a pattern of increasing stability over time, as predicted by the group-specific stereotypes theory.

Method

Subjects and procedure. The subjects in this study were introductory psychology students who agreed to participate in four 1-hr laboratory sessions in exchange for course credit. Prior to the first session, individuals were assigned to groups of 4 to 8 subjects each. In order to make the perceiver score estimates comparable across weeks (i.e., so that perceiver effects were always derived relative to a constant set of peers), the sample for the present study was restricted to those subjects who came to all four sessions and thus provided complete data. This final sample (*N* = 152; 51% female) represented 72% of the total number of subjects who attended the first week. The large majority of subjects were strangers at the start of the study: 97% of all possible pairings reported that they did not know one another at all, and fewer than 1% described their relationship as a preexisting friendship.

Subjects completed the measure of attachment styles prior to interacting with one another. The groups interacted for about 20 min the first week and about 40 min all subsequent weeks. The

group interactions were facilitated by an experimenter. During Week 1, subjects completed a "Lost on the Moon" group problem-solving task (see Robins & Beer, 2001). In Week 2, subjects completed a modified closeness-generating task in which they took turns answering questions that gradually increased in self-disclosure (adapted from Aron, Melinat, Aron, Vallone, & Bator, 1997). In Week 3, subjects had a leaderless group discussion in which they role-played members of an alumni awards committee; each individual had to both advocate for one candidate and work toward a group consensus (task adapted from John & Robins, 1994). In Week 4, the subjects played the party game *Beyond Balderdash*. At the end of each of the four sessions, after the group interaction, subjects were physically separated by the experimenter to complete ratings of interpersonal and self-perceptions.⁴

Measures.

Attachment style. Individual differences in attachment were measured using the four-paragraph Relationship Questionnaire (Bartholomew & Horowitz, 1991). In this measure, subjects read four paragraphs, each one describing a prototypical individual with a secure, preoccupied, fearful, or dismissing attachment style. Subjects rated how well each paragraph described them on a 7-point scale ranging from 1 (*disagree strongly*) to 7 (*agree strongly*). Thus, each attachment style was measured as a single item.

Interpersonal perceptions. Following each group interaction, other-reports and self-reports of the Big Five were collected with a single item for each personality trait. The items were as follows: for Extraversion, "This person is extraverted (energetic, assertive, sociable);" for Agreeableness, "This person is agreeable (warm, sympathetic, trusting);" for Conscientiousness, "This person is conscientious (responsible, thorough, efficient);" for Neuroticism, "This person is nervous (anxious, tense, moody);" and for Openness, "This person is open to new experiences (curious, imaginative, unconventional)." Each item was rated on a 10-point scale ranging from 0 (*disagree very strongly*) to 10 (*agree very strongly*). Perceiver effects were calculated in the same way as in Study 1, and again the perceiver effects and self-reports were group mean-centered. Perceiver effects were calculated separately for each item at each week. Intercorrelations of item means are reported in the Appendix.

Results and Discussion

Model comparisons. We compared a series of nested models specifying a single global evaluation factor, agency and communion factors, or the Big Five factors. These models were fit to the 20 perceiver effects in the study (5 trait dimensions each measured 4 times). Perceiver effects from different weeks served as multiple indicators for the substantive factors, so the factors all represented temporally stable variance. In addition to the substantive factors, all models included four within-week "method" factors with loadings from different trait perceptions within the same week; these method factors were fixed to be orthogonal to each other and to the substantive factors. The model comparisons led to the same conclusions whether or not these method factors were included, though absolute fit was better with the method factors. In contrast to Study 1, we only had one item per week for each of the trait perceptions, so we did not specify an acquiescence factor.

As with Study 1, we began with a model specifying method effects (i.e., the within-time-point factors) but no substantive factors. Unsurprisingly, this model showed poor fit; see Table 5. However, the primary purpose of this model was as a baseline on which to build more complex models with substantive factors.

The global evaluation model specified a single factor with loadings from all 20 perceiver effects. This model was a significant improvement over the method-only model, $\Delta\chi^2(20) = 238.6, p < .001$; and as shown in Table 5, all of the fit indices improved. All of the factor loadings were in the expected direction.

The third model included an agency factor, with loadings from the Extraversion and Openness perceiver effects, and a communion factor, with loadings from Agreeableness, Conscientiousness, and Neuroticism. This model fit significantly better than the global evaluation model, $\Delta\chi^2(1) = 47.9, p < .001$. Additionally, as shown in Table 5, all of the parsimony-adjusted fit statistics indicated that this model was preferred over the global evaluation model. All of the perceiver effects loaded in the expected direction on the agency and communion factors.

The fourth model specified factors for each of the Big Five. This model showed a significant improvement over the agency/communion model, $\Delta\chi^2(9) = 185.5, p < .001$. All of the fit indices indicated that this was the preferred model, even after adjusting for parsimony; see Table 5. Loadings were all in the expected direction. Absolute standardized loadings ranged from .37 to .87, and all were significant at $p < .001$.

As in Study 1, we noted that the latent Big Five factors had moderate to high intercorrelations (absolute factor correlations ranged from .31 to .81; see Table 3 below the diagonal). We therefore fit a model in which these intercorrelations were explained by a higher order global evaluation factor. As shown in Table 5, the chi-square was significantly worse, CFI was about the same, SRMR was worse, and BIC and RMSEA showed modest improvements; we thus concluded that the higher order factor was a reasonable fit to the factor intercorrelations. The weakest absolute loading of a latent trait factor on the higher order factor was $-.46$ for Neuroticism; the strongest was $.90$ for Conscientiousness. The specific factor variances for Extraversion, Agreeableness, Conscientiousness, and Neuroticism were significantly different from zero, indicating that the higher order global evaluation factor did not completely explain these trait factors. We also tested a model with two higher order factors (for agency and communion), but that model had worse fit on all indices.

In a follow-up analysis, we examined whether the structure of perceiver effects was the same in the first 2 weeks, when subjects did not know each other as well, and in the last 2 weeks, when subjects knew each other better. We fit one-, two-, and five-factor models separately in these subsets of the data. In both subsets of the data, the model comparisons indicated that the five-factor model fit better than the other models according to all of the fit indices.

Investigating model fit. In contrast to Study 1, the five-factor model in Study 2 showed acceptable absolute fit without any

⁴ Different findings from this data set have previously been reported by Srivastava and Beer (2005) and by Anderson, Srivastava, Beer, Spataro, and Chatman (2006). Neither report included analyses of the Big Five perceiver effects that are the focus of the present investigation.

Table 5
Confirmatory Factor Models of Perceiver Effects in Study 2

Model	χ^2	<i>df</i>	CFI	BIC	RMSEA	SRMR
1. Method factors only	685.9 ^a	170	.56	6445	.158	.285
2. Global evaluation plus method factors	447.3 ^b	150	.75	6302	.128	.091
3. Agency and communion plus method factors	399.4 ^c	149	.79	6259	.118	.089
4. Five factors plus method factors	213.9 ^d	140	.94	6117	.066	.058
5. Five factors plus method factors; trait factors load on a higher order global evaluation factor	217.6 ^e	145	.94	6096	.064	.060

Note. $N = 152$. Because the variables were group mean-centered, statistics were calculated with the sample size as $(N - g + 1)$, where g = the number of groups. Chi-square values with different superscripts are significantly different from each other at $p < .001$. CFI = comparative fit index; BIC = Bayesian information criterion; RMSEA = root-mean-square error of approximation; SRMR = standardized root-mean-square residual.

further modifications. To provide a basis for interpreting the absolute fit of the models for perceiver effects, we fit the same models to the self-perception data. The model comparisons clearly favored a five-factor model in the self-reports. The standardized fit statistics for the five-factor model of self-reports (with no higher order factors) were generally similar to those from the comparable model of perceiver effects: CFI = .99 (compared with .94 in the model of perceiver effects), RMSEA = .036 (compared with .066), and SRMR = .046 (compared with .058).

In summary, the model comparisons replicated the finding from Study 1 that a five-factor model of the perceiver effects is preferable over simpler models of global evaluation or agency and communion; some but not all of the trait variance could be explained by a higher order global evaluation factor. Furthermore, these perceiver effects were at least somewhat stable over the course of the longitudinal study.

Assumed similarity. As shown in Table 4, the zero-correlations between self-perceptions and perceiver effects showed evidence of assumed similarity. The magnitude of correlations at the start of the study was somewhat lower than in Study 1, though they were still generally within the range across multiple studies reported by Kenny (1994). As has been observed in previous studies, the assumed similarity correlations tended to increase over time (Kenny, 1994; Park & Judd, 1989).

Similar to Study 1, the partial correlations were generally somewhat smaller in magnitude than the zero-order correlations. The trait with the strongest pattern of trait-specific assumed similarity was Extraversion, which was significant 3 out of 4 weeks and grew to a partial correlation of .42 during Week 4. The trait with the weakest pattern was Agreeableness, which had a near-zero partial correlation for the first 3 weeks.

Generalized other and group stereotypes.

Individual differences. To examine relations between individual differences and perceiver effects, we averaged the perceiver effects across all 5 weeks and examined correlations with gender and attachment style (see Table 6). As hypothesized, gender was significantly correlated with perceiver effects for all five traits, such that women saw others as higher in Extraversion, Agreeableness, Conscientiousness, and Openness, and lower in Neuroticism. The direction of effects raised the possibility that the effect may have been at the level of global evaluations. To test this possibility, we calculated partial correlations between gender and each perceiver effect, controlling for the other perceiver effects. Partialing nearly eliminated the correlations of gender with Extraversion

(-.04), Agreeableness (.05), and Openness (.08), none of which was significant. The partial correlations were reduced but still significant at $p < .05$ with Conscientiousness (-.25) and Neuroticism (.16).

Attachment styles did not show evidence of a global evaluative bias. Just two significant correlations emerged: Securely attached individuals saw others as higher in extraversion, and preoccupied individuals saw others as less agreeable.

Stability of perceiver effects. If perceiver effects consolidate over time into group-specific stereotypes, then they should become increasingly stable over time as individuals form impressions of their groups. In other words, the correlation of Week 1 with Week 2 should be smaller than the correlation of Week 2 with Week 3, and that correlation should be smaller than the correlation of Week 3 with Week 4. Table 7 shows the 1-week stabilities of perceiver effects, and it can be seen that this was indeed the case: All of the perceiver effects showed a pattern of increasing stability over time, with the exception of the perceiver effect for openness. To test whether these increases were significant, we used Mplus to fit a model of the covariance matrix in which intervals with equal spacing were constrained to have equal covariances (i.e., a Toeplitz matrix). For example, in this model the covariance between Weeks 1 and 2 had to equal the covariance between Weeks 2 and 3. For all trait dimensions except Openness, the constrained model fit significantly worse than a model in which stability was free to increase over time. Averaging across traits, the stability corre-

Table 6
Correlations of Perceiver Effects With Gender and Attachment Style in Study 2

Perceiver effect	Gender	Attachment style			
		Secure	Preoccupied	Fearful	Dismissing
Extraversion	-.21	.20	-.05	-.10	-.07
Agreeableness	-.20	.09	-.21	-.04	.08
Conscientiousness	-.34	.09	-.06	-.06	.06
Neuroticism	.26	-.10	.10	-.03	-.03
Openness	-.18	.02	-.04	-.02	.06

Note. $N = 152$. Numbers in italics are significantly different from zero at $p < .05$. Gender was coded 0 = female, 1 = male. Perceiver effects are averaged across all 4 weeks.

Table 7
Stability of Perceiver Effects Over Time in Study 2

Perceiver effect	Correlation between weeks		
	1 and 2	2 and 3	3 and 4
Extraversion	.35	.51	.59
Agreeableness	.46	.50	.64
Conscientiousness	.29	.43	.56
Neuroticism	.42	.48	.52
Openness	.23	.59	.53
<i>M</i>	.37	.55	.65

Note. $N = 152$. All correlations are significantly different from zero, $p < .01$. Means were computed using Fisher r -to- z' transformation.

lations went from .37 between Weeks 1 and 2, to .55 between Weeks 2 and 3, to .65 between Weeks 3 and 4.⁵

Summary. In summary, in Study 2 we replicated the finding that a five-factor structure of perceiver effects is preferable to one- and two-factor models. The perceiver effects again showed evidence of assumed similarity, which was partially attributable to trait-specific variance. Perceiver effects were correlated with gender and attachment and were stable early in the study, consistent with a generalized other; but stability correlations increased over time, consistent with the formation of new, group-specific stereotypes.

General Discussion

The SRM has been an important paradigm for studying a wide range of phenomena in interpersonal perception, but perceiver effects may be one of its least studied aspects. Across two studies, we examined a series of questions in order to better understand perceiver effects in interpersonal perception. On the basis of model comparisons that replicated across studies, we found evidence that neither a single dimension of global evaluation nor two dimensions of agency and communion fit the perceiver effects as well as a model specifying five differentiated dimensions of meaning. Across both studies, there was evidence of a modest degree of assumed similarity: Individuals tended to see the self and others in similar ways, and some of this similarity appeared to be based on trait-specific content, not just broad evaluative tone. Perceiver effects were correlated with gender and attachment style, but they became increasingly stable over time; this suggested that perceiver effects were initially seeded from preexisting stable individual differences, but then they further consolidated over time as group-specific stereotypes.

Perceiver Effects Are Multidimensional

For a long time, the literature on biases in self-perception and interpersonal perception has assumed that such biases are organized around a single dimension of evaluation or desirability. Such a unidimensional model is implicit in the definition of the halo effect (Nisbett & Wilson, 1977), and it is explicit in many models of socially desirable responding (Block, 1965; Crowne & Marlow, 1960). Recent advances in the study of self-perception biases have indicated that the unidimensional model may be inad-

quate; instead, there are multiple dimensions along which the self may be distorted (Paulhus & John, 1998). The present results suggest that biases in perceptions of others also require taking a more differentiated view. Global evaluation is part of the story but not all of it.

A common practice in studies of social perception bias is to aggregate across multiple traits to produce an index of global evaluative bias. The present results suggest that in studies of such biases, simple aggregation may blur important trait-specific content. This conclusion is compatible with a trend in stereotyping research toward trying to understand the different social and affective antecedents and consequences of stereotypes with different content (Fiske et al., 2002). This approach may also apply to studies of self-perception, especially in light of the finding that assumed similarity often depends at least partially on trait-specific content. If perceiver effects are better described by a five-factor model than by a two-factor model, then it may be important to revisit the structure of self-perception biases to see whether trait-specific distortions can be meaningfully differentiated from broad concerns about agency and communion (Paulhus & John, 1998).

Origins and Explanations of the Perceiver Effect

Where do perceiver effects come from? Longer term longitudinal research will be important to better understand the early origins of perceiver effects, but the present results provide some important clues. Robust correlations with gender, and some modest correlations with attachment style, suggest that adult perceivers come to new social groups with preexisting expectations and stereotypes that guide their initial impressions of others. The implication is that if one were to compute perceiver effects for the same individuals participating in more than one social group, some portion of the perceiver variance would be consistent across groups. However, the pattern of increasing stability over time found in this study suggests that perceiver effects are not brought fully formed (and fully stable) into a new social context. Rather, as individuals interact and get to know fellow group members, their initial expectations consolidate into a group-specific stereotype, which may account for an additional (and increasing) portion of the variance in perceiver effects. The solidification of these stereotypes may also help explain why overall variance in perceiver effects tends to increase over time.

Should perceiver effects be interpreted as a "bias" in interpersonal perception, and if so, where do these biases come from? A common definition of bias in interpersonal perception is the difference between a perception and some objective criterion or standard of accuracy. If Iris, Jack, Ken, and Laura all observe one another simultaneously in a group setting, and if we expect that trait ratings can and should reflect an objective reality, then the perceiver effect may be interpreted as an index of relative bias in trait judgment: if $P_{Iris} > P_{Jack}$, then we can conclude that Iris's

⁵ This pattern of increasing stability held true even when we calculated disattenuated stability correlations using the reliability coefficients reported in Table 1; thus, the pattern did not appear to be an artifact of changes in the reliability of the perceiver effects. The disattenuated stability correlations, averaged across traits, increased from .49 to .74.

bias (the difference between her perceptions and some objective standard of accuracy) is more positive than Jack's bias, even if we do not know the value of the objective standard.

However, when we consider the social cognitive and social behavioral processes involved in interpersonal perception, the term *bias* might be problematic, or at least oversimplified. One possibility is that Iris and Jack might differ in some important social cognitive process involved in person perception. Kenny (1994) proposed that self-perceiver correlations (assumed similarity) result from common enhancement processes that are applied to self-perceptions and to perceiver effects (see also Kwan et al., 2004). The finding in the present study that assumed similarity has trait-specific content suggests that there are multiple such enhancement processes that account for different perceiver effects.

Another possibility is that group members might behave differently toward Iris than they do toward Jack (perhaps because of appearance-based cues like gender, or elicited by Iris's and Jack's different behaviors; see Kenny, 1994). If Iris is treated more kindly than Jack and both rely more heavily on their personal experience in forming trait perceptions, then the result may be that Iris attributes more agreeableness to others than Jack does because Iris has observed more agreeable behavior directed toward herself than Jack has observed directed toward himself. In this case, the operational definition of "biases" still holds—Iris and Jack still differ in how much they depart from an objective standard in rating others. But the reason for Iris's and Jack's different perceptions is rooted in their different social experiences, not in differences between Iris's and Jack's cognitive processing.

These two explanations for the origins of bias are important in interpreting the correlations between individual differences and perceiver effects. For example, we found that women perceived others more positively than did men. Gender-based socialization may instill in women a sense of relatedness and connection with others, and in men a tendency to view others as separate from the self or even as competitors (Cross & Madson, 1997); this may give rise to cognitive biases in perceptions of others. However, it is important to also consider that many stereotypes of women are evaluatively favorable: Women are often viewed by others through a paternalistic lens as being incapable but warm (Eagly & Mladinic, 1989; Fiske et al., 2002). If a result of these stereotypes is that women are frequently treated by others in a more superficially pleasant way, then women may attribute more positive traits to others based on their social experiences.

Implications for Personality Assessment

The present research has implications for designing multimethod assessments and for analyzing their results. It is well established that inventory reports based on self-perceptions and peer perceptions of Big Five traits are valid predictors of behavior and meaningful life outcomes (e.g., Ozer & Benet-Martínez, 2006; Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), and in general, self and peers agree substantially in their perceptions of personality (Funder, 1980; Watson, Hubbard, & Weise, 2000). At the same time, nonzero and even nontrivial validity does not rule out the possibility of some bias in measurements, and an integrated model of interpersonal perception must account for both valid judgment and for perceptual processes that can go awry (Funder, 1995).

Psychologists have been long aware of the possibility of various biases in self-reports and have devoted considerable attention to understanding them (e.g., Block, 1965; Paulhus, 2002). For this reason, other-reports from peers, romantic partners, parents, and other informants have been proposed as a useful way of measuring personality. Some recommendations have favored other-reports over self-reports (Hofstee, 1994), whereas others have proposed that both self- and other-reports be used as part of an overall multimethod assessment strategy (Kraemer et al., 2003; Vazire, 2006). The present results, which document stable and differentiated biases in other-reports, suggest that no single data source is free of bias and that multimethod triangulation is likely the best approach.

The present findings also open up new avenues for research into the best ways to use multimethod assessments. Most studies that use informant reports draw upon friends or other prior acquaintances of the research subjects, and researchers often ask their subjects to select the informants. It will therefore be important to study how perceiver effects are distributed in social networks. If multiple informants recruited by a single subject are likely to share biases with one another, or to have biases that are congruent with the self-perception biases of the subject who recruited them (perhaps due to social calibration by the research subject), then multi-informant designs will still be subject to systematic biases that will need to be dealt with in some other way. The present results offer some guidance for such investigations: First, studies of perceiver biases among informants will need to use multidimensional measurement models, rather than relying on global evaluation or related constructs like social desirability. Second, researchers also cannot assume that perceiver effects are highly consistent across contexts, as one would expect from the generalized-other theory. Rather, informants' perceiver effects will need to be measured in contexts and social groups that are relevant to the informant's relationship with the subject of interest.

Implications for Studying Interpersonal Functioning and Psychopathology

Perceiver effects may be relevant to understanding interpersonal difficulties associated with psychopathology. Research on narcissism has documented systematic distortions in narcissists' self-perceptions (John & Robins, 1994), and analogous biases in perceiver effects may be an important part of the cognitive framework underpinning other forms of psychopathology. For example, distorted perceptions of others have been associated with both borderline personality disorder (Benjamin & Wonderlich, 1994), depressive disorder (Gara et al., 1993), and conduct disorder (Dodge, 1993). The finding that perceiver effects have a multidimensional structure may provide a framework for organizing this literature. For example, adolescents with conduct disorder tend to attribute hostile traits (i.e., low agreeableness) to peers; the present results suggest that we should not necessarily assume that they make other negative attributions.

At a first approximation, it is possible that certain disorders may be associated with maladaptive conceptions of the generalized other. A more nuanced approach, and one that accounts for the finding that perceiver effects are context dependent, would be to compare perceiver effects formed in relation to different sets of others across different contexts. For example, some research has

suggested that depressed individuals might be more likely to use social projection, and therefore to form more negative perceptions of others, in contexts in which they feel close to or dependent on others (Gara et al., 1993). Similarly, the correlations between attachment style and perceiver effects might have been stronger in a context in which attachment concerns were more relevant and the attachment system was more highly activated, such as in groups of people who expected to rely on one another for acceptance and support.

Limitations and Future Directions

The present studies had a number of limitations that highlight opportunities for future research. All of the social interactions involved groups with a relatively friendly pretext for being together (although the activities in Study 2 did include a group cognitive task and a leaderless group discussion). Similarity in contexts helps us to make comparisons across studies, but it will be important in future studies to examine other kinds of social interactions, such as groups engaged in overtly competitive activities. Different activities might produce different results. For example, Ames (2004) found that individuals are more likely to use social projection with others that they believe are generally similar to the self and more likely to use stereotypes with others that are believed to be generally dissimilar. It is possible that in a social context in which subjects did not feel particularly connected to others, they may have relied more heavily on preexisting stereotypes—perhaps producing clearer generalized-other effects as a result.

In Study 2, the increasing stability over time was consistent with a local, group-specific stereotype that was being molded over time. A more direct test, however, would have been to have subjects each interacting in multiple different groups and to compute perceiver effects independently in different groups. Although such a procedure would be much more involved, it opens up a number of interesting possibilities, such as varying the makeup of different groups or their pretext for interacting (e.g., for communal or exchange purposes; Clark & Mills, 1979).

It would also be worth exploring these questions in other populations. As already noted, perceiver effects may have interesting applications to understanding clinical disorders. Perceiver effects may also be sensitive to developmental changes and cultural differences. Perceiver effects are an understudied component of interpersonal perception, and our goal was to help build a foundation for future substantive research. We hope that the present results will open up new possibilities for studying perceiver effects in relation to a range of issues.

References

- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55*, 387–395.
- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*, 256–274.
- Ames, D. R. (2004). Strategies for social inference: A similarity contingency model of projection and stereotyping in attribute prevalence estimates. *Journal of Personality and Social Psychology, 87*, 573–585.
- Anderson, C., Srivastava, S., Beer, J. S., Spataro, S. E., & Chatman, J. E. (2006). Knowing your place: Self-perceptions of status in social groups. *Journal of Personality and Social Psychology, 91*, 1094–1110.
- Aron, A., Melinat, E., Aron, E. N., Vallone, R., & Bator, R. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin, 23*, 363–377.
- Bakan, D. (1966). *The duality of human existence: Isolation and communion in Western man*. Boston: Beacon Press.
- Baldwin, M. W., Keelan, J. P. R., Fehr, B., Enns, V., & Koh-Rangarajoo, E. (1996). Social cognitive conceptualization of attachment styles: Availability and accessibility effects. *Journal of Personality and Social Psychology, 71*, 94–109.
- Bartholomew, K., & Horowitz, L. M. (1991). Attachment styles among young adults: A test of a four-category model. *Journal of Personality and Social Psychology, 61*, 226–244.
- Benjamin, L. S., & Wonderlich, S. A. (1994). Social perceptions and borderline personality disorder: The relation to mood disorders. *Journal of Abnormal Psychology, 103*, 610–624.
- Biesanz, J. C., & West, S. G. (2004). Towards understanding assessments of the Big Five: Multitrait-multimethod analyses of convergent and discriminant validity across measurement occasion and type of observer. *Journal of Personality, 72*, 845–876.
- Block, J. (1965). *The challenge of response sets*. New York: Appleton-Century-Crofts.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Beverly Hills, CA: Sage.
- Chen, S., & Andersen, S. M. (1999). Relationships from the past in the present: Significant-other representations and transference in interpersonal life. *Advances in Experimental Social Psychology, 31*, 123–190.
- Clark, M. S., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology, 37*, 12–24.
- Cronbach, L. J. (1955). Processes affecting scores on “understanding others” and “assumed similarity.” *Psychological Bulletin, 52*, 177–193.
- Cross, S. E., & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin, 122*, 5–37.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349–354.
- DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology, 91*, 1138–1151.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417–440.
- Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246–1256.
- Dodge, K. A. (1993). Social-cognitive mechanisms in the development of conduct disorder and depression. *Annual Review of Psychology, 44*, 559–584.
- Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin, 15*, 543–558.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*, 117–140.
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878–902.
- Foa, U. G., & Foa, E. B. (1974). *Societal structures of the mind*. Springfield, IL: Charles C. Thomas.
- Fraley, R. C. (2002). Attachment stability from infancy to adulthood: Meta-analysis and dynamic modeling of developmental mechanisms. *Personality and Social Psychology Review, 6*, 123–151.

- Funder, D. C. (1980). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *Journal of Personality*, 48, 473–493.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102, 652–670.
- Gara, M. A., Woolfolk, R. L., Cohen, B. D., Goldston, R. B., Allen, L. A., & Novalany, J. (1993). Perception of self and other in major depression. *Journal of Abnormal Psychology*, 102, 93–100.
- Goldberg, L. R. (1990). An alternative “description of personality”: The Big-Five factor structure. *Journal of Personality and Social Psychology*, 59, 1216–1229.
- Gosling, S. D. (2008). *Snoop: What your stuff says about you*. New York: Basic Books.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the Big Five personality domains. *Journal of Research in Personality*, 37, 504–528.
- Hofstee, W. K. B. (1994). Who should own the definition of personality? *European Journal of Personality*, 8, 149–162.
- Hogan, R. (1983). A socioanalytic theory of personality. In M. M. Page (Ed.), *Nebraska symposium on motivation, 1982* (pp. 55–89). Lincoln: University of Nebraska Press.
- John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, 61, 521–551.
- John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology*, 66, 206–219.
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York: Guilford Press.
- Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York: Guilford Press.
- Kenny, D. A. (2004). PERSON: A general model of interpersonal perception. *Personality and Social Psychology Review*, 8, 265–280.
- Kenny, D. A., & La Voie, L. (1984). The social relations model. *Advances in Experimental Social Psychology*, 18, 141–182.
- Kenny, D. A., Lord, R., & Garg, S. (1986). *A social relations analysis of peer ratings*. Unpublished manuscript, University of Connecticut.
- Kraemer, H. C., Measelle, J. R., Ablow, J. C., Essex, M. J., Boyce, T., & Kupfer, D. J. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *American Journal of Psychiatry*, 160, 1566–1577.
- 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, 94–111.
- MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, 111, 490–504.
- Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: An integrative hierarchical approach. *Journal of Personality and Social Psychology*, 88, 139–157.
- McAdams, D. P. (1985). *Power, intimacy, and the life story: Personological inquiries into identity*. New York: Guilford Press.
- McCrae, R. R., Zonderman, A. B., Costa, P. T., Jr., & Bond, M. H. (1996). Evaluating replicability of factors in the Revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, 70, 552–566.
- Mead, G. H. (1934). *Mind, self, and society*. Chicago: University of Chicago Press.
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide* (4th ed.). Los Angeles, CA: Author.
- Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, 35, 250–256.
- Ozer, D., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, 57, 401–421.
- Park, B., & Judd, C. M. (1989). Agreement on initial impressions: Differences due to perceivers, trait dimensions, and target behaviors. *Journal of Personality and Social Psychology*, 56, 493–505.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 67–88). Hillsdale, NJ: Erlbaum.
- Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic bias in self-perceptions: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, 66, 1025–1060.
- Roberts, B. W., Kuncel, N., Shiner, R. N., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socio-economic status, and cognitive ability for predicting important life outcomes. *Perspectives in Psychological Science*, 2, 313–345.
- Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: Short-term benefits and long-term costs. *Journal of Personality and Social Psychology*, 80, 340–352.
- Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality*, 69, 847–880.
- Saucier, G., Ostendorf, F., & Peabody, D. (2001). The non-evaluative circumplex of personality adjectives. *Journal of Personality*, 69, 537–582.
- Smith, E. R. (1998). Mental representation and memory. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 1, pp. 391–445). New York: McGraw-Hill.
- Srivastava, S., & Beer, J. S. (2005). How self-evaluations relate to being liked by others: Integrating sociometer and attachment perspectives. *Journal of Personality and Social Psychology*, 89, 966–977.
- Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality*, 40, 472–481.
- Warner, R., Kenny, D. A., & Stoto, M. (1979). A new round robin analysis of variance for social interaction data. *Journal of Personality and Social Psychology*, 37, 1742–1757.
- Watson, D., Hubbard, B., & Weise, D. (2000). Self–other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, 78, 546–558.
- Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In W. M. Grove & D. Cicchetti (Eds.), *Thinking clearly about psychology: Vol. 2. Personality and psychopathology* (pp. 89–113). Minneapolis: University of Minnesota Press.
- Winquist, L. A., Mohr, C. D., & Kenny, D. A. (1998). The female positivity effect in the perception of others. *Journal of Research of Personality*, 32, 370–388.

(Appendix follows)

Appendix

Table A1
Correlations Among Single-Item Self-Reports and Perceiver Effects in Study 1

Item	E	E (R)	A	A (R)	C	C (R)	N	N (R)	O	O (R)
Extraversion	—	.51	.27	-.04	.14	.03	-.06	-.11	.33	.15
Extraversion (R)	.31	—	.02	-.03	-.05	-.05	-.15	.08	.22	.19
Agreeableness	.19	-.02	—	.17	.26	.25	-.11	-.20	.14	.20
Agreeableness (R)	.07	.17	.14	—	.00	.19	-.41	-.17	-.01	.09
Conscientiousness	.19	.01	.40	.06	—	.35	-.03	-.31	.21	.11
Conscientiousness (R)	.08	.08	.25	.29	.28	—	-.11	-.22	-.03	.13
Neuroticism	-.15	-.30	-.20	-.47	-.23	-.36	—	.41	-.14	-.12
Neuroticism (R)	-.01	.12	-.24	-.22	-.33	-.19	.31	—	-.23	-.00
Openness	.37	.21	.21	.08	.32	.04	-.24	-.10	—	.27
Openness (R)	.16	.20	.26	.29	.15	.37	-.43	-.10	.19	—

Note. $N = 423$. Self-reports are above the diagonal; perceiver effects are below the diagonal. R = reverse-keyed item (these items were reversed prior to computing correlations).

Table A2
Correlations Among Average Self-Reports and Perceiver Effects in Study 2

Item	1	2	3	4	5
1. E	—	.52	.31	-.45	.52
2. A	.64	—	.61	-.16	.43
3. C	.62	.69	—	-.16	.33
4. N	-.24	-.29	-.36	—	-.21
5. O	.62	.65	.69	-.28	—

Note. $N = 152$. Self-reports are above the diagonal; perceiver effects are below the diagonal. E = Extraversion; A = Agreeableness; C = Conscientiousness; N = Neuroticism; O = Openness.

Received December 4, 2008
Revision received May 28, 2009
Accepted July 2, 2009 ■