American Psychological Association

Journal of Personality and Social Psychology:
Personality Processes and Individual Differences

ISSN: 0022-3514

2021, Vol. 120, No. 3, 745–764
http://dx.doi.org/10.1037/pspp0000278

PERSONALITY PROCESSES AND INDIVIDUAL DIFFERENCES

What Is the Structure of Perceiver Effects? On the Importance of Global Positivity and Trait-Specificity Across Personality Domains and Judgment Contexts

Richard Rau
University of Münster

Erika N. Carlson
University of Toronto Mississauga

Mitja D. Back
University of Münster

Maxwell Barranti
University of Toronto Mississauga

Jochen E. Gebauer
University of Mannheim and University of Copenhagen

Lauren J. Human
McGill University

Daniel Leising
Technical University of Dresden

Steffen Nestler
University of Münster

When judging others' personalities, perceivers differ in their general judgment tendencies. These perceiver effects partly reflect a response bias but are also stable and psychologically important individual differences. However, current insights into the basic structure of perceiver effects are ambiguous with previous research pointing to either a unidimensional structure (i.e., people see others as globally positive vs. negative) or a multidimensional structure (i.e., people see others as high or low on specific traits). Here we provide a large scale investigation of the structure of perceiver effects that spans more than 100,000 personality judgments across 10 studies in which a total of $N = 2{,}199$ perceivers judged others on several trait domains (i.e., the Big Five, agency & communion) and in different judgment contexts (i.e., level of involvement with targets, level of exposure to targets). Results suggest that perceiver effects are hierarchically structured such that they reflect both a global tendency to view others positively versus negativity and specific tendencies to view others as high or low with respect to trait content. The relative importance of these components varied considerably across trait domains and judgment contexts: Perceiver effects were more specific for traits higher in observability and lower in evaluativeness and in context with less personal involvement and higher exposure to targets. Overall, results provide strong evidence for the hierarchical structure of perceiver effects and suggest that their

This article was published Online First December 12, 2019.

Richard Rau, Department of Psychology, University of Münster; Erika N. Carlson, Department of Psychology, University of Toronto Mississauga; Mitja D. Back, Department of Psychology, University of Münster; Maxwell Barranti, Department of Psychology, University of Toronto Mississauga; Jochen E. Gebauer, Department of Psychology, University of Mannheim, and Department of Psychology, University of Copenhagen; Lauren J. Human, Department of Psychology, McGill University; Daniel Leising, Department of Psychology, Technical University of Dresden; Steffen Nestler, Department of Psychology, University of Münster.

Maxwell Barranti is now at the Department of Psychology, York University.

Additional materials including data, R-Code, supplemental analyses, and a list of prior publications using data presented here are retrievable from osf.io/kr5ms/. An overview of all additional materials is presented in Appendix C.

We thank Lisa M. Niemeyer and Simine Vazire for sharing data, Florian Scharf and Stefan Schmukle for statistical advice, and David A. Kenny for feedback on an earlier draft of this article. The preparation of this article was supported in part by a grant from the German Research Foundation to Mitja D. Back (BA 3731/8-1), a grant from the Social Sciences and Humanities Research Council to Erika N. Carlson (435150611), a grant from the German Research Foundation to Jochen E. Gebauer (GE 2515/6-1), a grant from the John Templeton Foundation to Matthias Mehl (44245), a grant from the German Research Foundation to Steffen Nestler (NE 1485/6-1), and a grant from the National Science Foundation to Simine Vazire (BCS-1025330).

Correspondence concerning this article should be addressed to Richard Rau, Department of Psychology, University of Münster, Fliednerstraße 21, Institute for Psychology, 48149 Münster, Germany. E-mail: richard.rau@uni-muenster.de

meaning systematically varies depending on trait domain and possibly the judgment context. Implications for theory and assessment are discussed.

Keywords: generalized other-perception, interpersonal perception, perceiver effect, person impression, person judgment

Arguably, we have all met the kind of person who has a rosy view of their social environment and who appears to see the best in virtually everybody he or she meets. In contrast, some individuals seem to have harsh views about everyone around them and seem to habitually find fault with others. Indeed, people can have strikingly different ways of looking at their social environments, and their idiosyncratic judgment tendencies might ultimately reflect general individual differences of perceivers. This paper aims to better understand these judgment tendencies by examining whether they exist only on a global level (indicating a generally rosy vs. harsh view) or also on the level of specific trait content such that, for instance, they indicate individual differences to see others as reserved versus sociable, simple versus creative, or careless versus dependable.

In the interpersonal perception literature, two perceivers who judge the same trait in the same set of target people but arrive at different average judgments are said to have different *perceiver effects* (Kenny, 1994). In the present research, we examined whether perceiver effects reflect only globally evaluative or also trait-specific tendencies in 10 studies that involved 2,199 perceivers who judged multiple targets on multiple personality dimensions. These judgments were made either on the basis of physical appearance, unstructured chats, icebreaking games, discussions of moral dilemmas, problem solving tasks, economic decision making games, creativity tasks, or group work on graded class assignments. Specifically, we had three goals: First, we aimed to identify the configural structure of perceiver effects, that is, we sought to learn about the existence of response bias, global positivity, and trait-specificity. Second, we quantified the relative importance of these components. Third, we examined if the relative importance of positivity and specificity depends on what trait is being judged and/or on the judgment context.

Perceiver Effects as an Important Individual Difference Variable

The idea that people have idiosyncratic ways of viewing others has a long track record in psychological theory. For instance, according to Erikson's stages of psychosocial development, individuals are thought to differentially develop *basic trust* toward others during infancy, and this sense of trust influences how individuals approach other people over the course of their lives (Erikson, 1959, 1968). Similarly, attachment theory posits that children internalize prototypical expectations about others which shape future relationships (Bowlby, 1988; Fraley, 2002). According to other popular accounts, generalized other-perceptions result from a tendency to assume that others are either similar to or different from the self (Campbell, Miller, Lubetsky, & O'Connell, 1964; Cronbach, 1955). Further, a negatively biased view of one's environment is thought to be a core aspect of depression (Beck, 1979) and generalized beliefs about others are at the heart of many personality disorders (Hopwood, Schade, Krueger, Wright, & Markon, 2013). As such, idiosyncratic perceptions of others are of high theoretical interest and resemble a pervasive individual difference variable that may crucially shape interpersonal processes and life outcomes.

Despite their long track record in psychological *theory*, idiosyncratic perceptions of others have a much shorter empirical history. In fact, perceiver effects were often conceptualized as a potential source of error (i.e., low interjudge agreement) in the endeavor of measuring characteristics of the target (e.g., people's reputations). This line of work has revealed that perceiver effects are a ubiquitous component of person perception which introduces variance (i.e., perceiver variance) to personality judgments for almost any trait and in almost any judgment context. With respect to traits, perceiver variance is well documented for all of the Big Five domains (for an overview, see Kenny, 1994) but also for many other traits outside of this taxonomy (Dufner, Leising, & Gebauer, 2016; Wood, Harms, & Vazire, 2010). With respect to contexts, substantial perceiver variance has been found in judgments of photographs of strangers (Hehman, Sutherland, Flake, & Slepian, 2017), judgments of members in newly formed social groups (Rau, Nestler, Geukes, Back, & Dufner, 2019), and judgments of close friends (Vazire, 2010). On average, perceiver effects account for approximately 20% to 30% of the variance in personality judgments (Hehman et al., 2017; Kenny, 1994). Perceiver effects are also fairly constant over time (Srivastava, Guglielmo, & Beer, 2010; Wood, Harms, et al., 2010), suggesting they might be a stable dispositional characteristic of the perceiver.

Importantly, personality judgments are typically examined within some kind of framework (e.g., the Big Five), but assumptions about the factor structures underlying such frameworks have rarely been tested for perceiver effects. However, basic knowledge about the structure of perceiver effects is crucial for their conceptual understanding: Beyond response bias, do perceiver effects reflect individual differences in general positivity and/or trait-specific judgment tendencies? Considering that perceiver effects are involved in many consequential decisions in everyday life (e.g., teacher evaluations, personnel selection, and clinical diagnoses), and given the theoretical importance of perceiver effects in many psychological disciplines, much value lies in increasing our knowledge about their structure.

Potential Factor Structures of Perceiver Effects

Typically, a perceiver effect is indexed by the average rating a perceiver provides about several target people on a given rating dimension. Let us, for example, assume that Peter judges the sociability of his classmates with an average rating of 9 on a scale from 1 to 10. For illustrative purposes, one can then subtract the average sociability rating across all perceivers, say $M = 7$, from Peter's score to find that his perceiver effect for sociability is +2 (but note that subtracting the grand mean is not required statistically). Theoretically, Peter's positive perceiver effect for sociabil-

ity could be attributable to four reasons: Measurement error, acquiescence bias, a global tendency toward positive judgments, and a specific tendency toward perceiving others as highly sociable. The first two reasons are not psychologically informative. Measurement error would signal a lack of reliability and acquiescence would signal Peter's tendency to use higher rather than lower numbers irrespective of what he rates. Yet, an interesting question remains open in the absence of measurement error and acquiescence. Does Peter's high perceiver effect for sociability result from a generally rosy view he holds about others or does it reveal a specific tendency toward perceiving high sociability in others? Given that sociability has a positive valence, it is impossible to disentangle these two potential sources without further information. To achieve such disentangling, Peter's perceiver effect on another dimension related to sociability (e.g., outgoingness) and on a dimension unrelated to sociability (e.g., creativity) is needed. Once this information is available, the question of positivity versus specificity can be formalized as different factor structures, which are displayed in Figure 1. On the one hand, if perceiver effects reflected the degree to which people see others positively or negatively across personality traits, this would be best represented by a single positivity factor (see Figure 1a). On the other hand, if most perceivers tended to see others as high on some traits and low on other traits, this would best be represented by trait-specific factors (see Figure 1b). Of course, it is also possible that perceiver effects reflect a blend of both global and trait-specific tendencies (see Figure 1c).

What is the conceptual meaning of a model that features both a positivity factor and trait-specific factors? The model implies that individuals' perceptions of other people are colored by their global tendency to see others in a positive or negative light but that a specific tone is added to these perceptions owing to their idiosyncratic views concerning particular trait content. For example, Peter might have a high perceiver effect on many positively valenced traits (he rates others' sociability, creativity, and relaxedness mostly with 8s, 9s, and 10s) but he might have a rather low perceiver effect for a single other trait (his average rating of others' dependableness is 5). If such patterns were common across perceivers, this would best be captured by the positivity-specificity model as opposed to the positivity-only or the trait-specificity model. The positivity factor would probably track individual differences in evaluative attitudes, that is, in how much perceivers generally *like* other people and the specific factors would indicate the personal issues perceivers have in a particular trait domain, irrespective of their evaluative attitude toward others. This would raise a number of psychologically interesting questions. For instance, what are the antecedents of perceiver effects (e.g., where did Peter's issue with dependableness come from?), how consistent are perceiver effects (e.g., are Peter's idiosyncratic perceptions the same at work as they are with friends?) and what are the interpersonal consequences of perceiver effects (e.g., how does Peter's generally positive attitude toward others affect his social relationships?).

Previous Findings on the Structure of Perceiver Effects

There have been two seminal investigations on the structure of perceiver effects to date. Srivastava et al. (2010) compared the three models depicted in Figure 1 in perceiver effects of Big Five ratings in two low acquaintance group interaction studies. Their approach was based on the observation that despite tapping into distinct descriptive content, all of the Big Five traits have a more-desirable and a less-desirable pole (John & Robins, 1993; Saucier, Ostendorf, & Peabody, 2001). Using confirmatory factor analysis, they found that the positivity-specificity model offered a better account of their data than both the positivity-only and the trait-specificity model. Further, perceiver effects partly reflected acquiescence bias. In the other investigation, Wood, Harms, et al. (2010) found evidence for a positivity-only model using principal

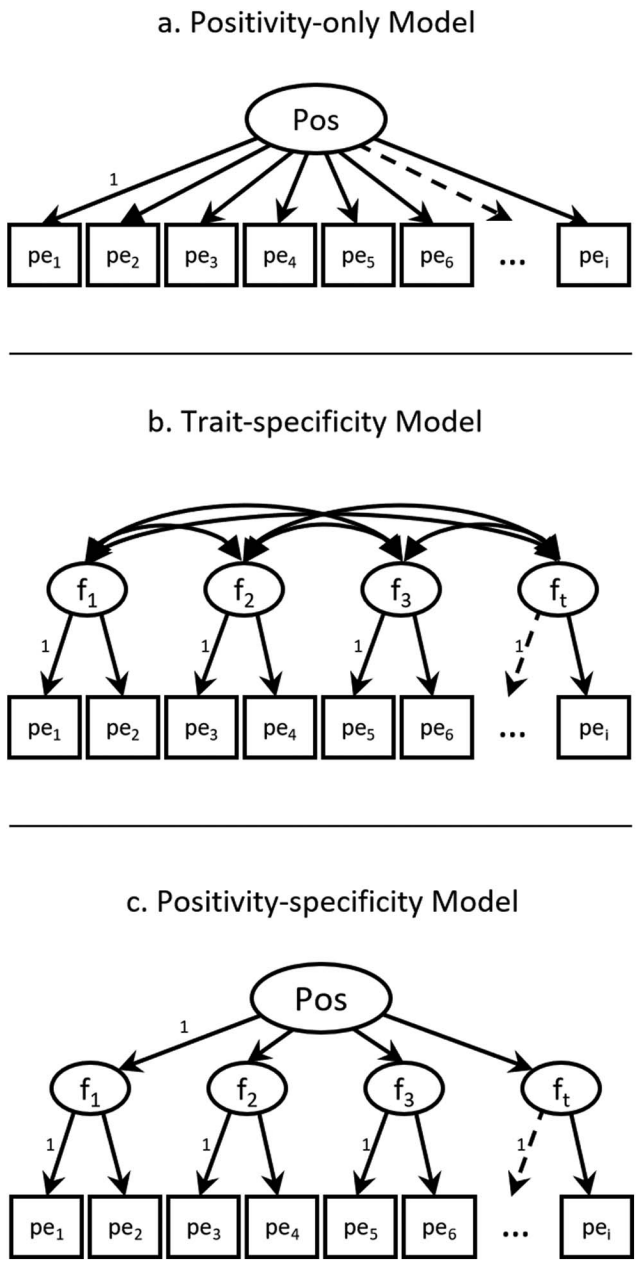

Figure 1. Potential factor models for perceiver effects of *i* items relating to *t* traits.

axis factoring in two studies involving personality ratings among roommates. Further, they also found evidence for acquiescence bias. Thus, the investigations converge with respect to acquiescence and global positivity but their results are discrepant with respect to the existence of trait-specificity. It is difficult to explain this discrepancy because the investigations differed in many methodological aspects (e.g., different trait measures, different factor analytical techniques) and many contextual aspects (e.g., different levels of involvement, different levels of exposure). Thus, considerable uncertainty about the structure of perceiver effects remains a hindrance for important research concerning the antecedents, consistency, and consequences of idiosyncratic perceptions of others. Therefore, the first goal of the present research was to clarify if a robust factor structure in perceiver effects exists.

The Relative Importance of Global Positivity and Trait-Specificity

Once a factor structure is established, an important follow-up question is, what are the relative contributions of the involved factors? If, for instance, a unidimensional model fits best, then a single factor can account for most of the covariation between observed variables. However, this says nothing about the *importance* of the factor because the factor might account for 10% or for 90% of the total variation and still fit well. Similarly, if a positivity-specificity structure was established for perceiver effects, this would indicate that both kinds of factors are in principle required, but it would still be important to learn about their relative contributions. For instance, if the general positivity factor accounted for substantially more variance than trait-specific factors, this would challenge the view that perceiver effects are projections of the perceivers' own traits (Campbell et al., 1964; Cronbach, 1955). If, on the other hand, trait-specificity accounted for substantially more variance than general positivity, this would challenge the view that people have highly generalized working models of others (Bowlby, 1988; Erikson, 1959, 1968). Given that previous research has focused exclusively on the configural structure of perceiver effects, the second goal of the present research was to explicitly quantify the contributions of positivity and specificity for the first time.

The Potential Influence of Trait Domain and Judgment Context on the Relative Importance of Global Positivity and Trait-Specificity

Importantly, the contributions of positivity and specificity are not necessarily constant across different trait domains and judgment contexts. Previous research has, however, not yet investigated how the role of positivity or specificity is impacted by traits and contexts. The third goal of the present research was therefore to explore whether the contributions of positivity and specificity depend on which traits are being judged and on the judgment context.

Do Contributions of Positivity and Specificity Depend on the Traits Being Judged?

In principle, perceiver effects can pertain to any personality content and to any personality framework. As outlined above, they are well documented for a large variety of traits. Previous research on the structure of perceiver effects has either attended to the Big Five framework (Srivastava et al., 2010) or applied large sets of personality items (Wood, Harms, et al., 2010). However, no prior work has considered perceiver effects within the agency-communion framework (the Big Two), which represent basic dimensions of interpersonal perceptions more generally. Ample evidence suggests that initial personality judgments are often intuitively made on these dimensions whereby agency taps into people's ability to promote the self (e.g., assertiveness, confidence) and communion taps into people's drive to establish social ties with others (e.g., warmth, cooperativeness; Abele & Wojciszke, 2007; Bakan, 1966; Leary, 1957; Wiggins, 1979).[1]

Of course, which framework researchers employ could impact the factor structures found for perceiver effects. Evidence for trait-specificity was only found by Srivastava et al. (2010) who used the Ten-Item Personality Inventory (TIPI; Gosling, Rentfrow, & Swann, 2003) and not by Wood, Harms, et al. (2010; Study 2) who used an early version of the Inventory of Individual Differences in the Lexicon (IIDL; Wood, Nye, & Saucier, 2010) to assess personality perceptions. The TIPI is a short measure specifically designed to capture the Big Five, whereas the IIDL is a 57-item measure designed to comprehensively cover the most frequently used trait terms without adhering to an overarching dimensional structure. Thus, it is currently unclear how strongly structure solutions for perceiver effects depend on the trait framework.

In the present research, we analyzed judgments of agreeableness, conscientiousness, extraversion, emotional stability, and openness as well as judgments of agency and communion, thus we considered both a Big Five and a Big Two framework. Importantly, these traits are not necessarily influenced by positivity and specificity to the same extent. As such, a major goal of the present research was to quantify the relative importance of positivity and specificity for each of these traits and explore whether perceiver effects reflect different tendencies across these traits.

Do Contributions of Positivity and Specificity Depend on the Judgment Context?

We have outlined above that perceiver effects can be observed quite ubiquitously across different judgment contexts. Notably, these contexts might fundamentally change the meaning of perceiver effects. For example, perceiver effects may be self-related ("how others are with me") when the perceiver is personally involved with targets but more generalized ("how others behave in general") when personal involvement is low (Kenny, 1994, p. 91). Further, it has been proposed that perceiver effects at low levels of exposure to targets (e.g., in first impressions of strangers) reflect a

[1] Srivastava et al. (2010) explored a metatrait model featuring two higher-order factors, one capturing shared variance among agreeableness, conscientiousness, and emotional stability and one capturing shared variance among extraversion and openness. Although there is little consensus concerning the labels of these meta-traits—suggestions include *alpha* and *beta* (Digman, 1997), *social propriety* and *dynamism* (Saucier & Goldberg, 2001), and *stability* and *plasticity* (DeYoung, 2006)—Srivastava et al. (2010) chose *agency* and *communion* as labels. However, in accordance with interpersonal circumplex theory, we use the labels agency and communion in a conceptually narrower sense herein.

stereotypical *expectation* about others, but that they reflect more of an *opinion* at higher levels of exposure to targets (e.g., in well acquainted groups; Kenny, 2004). As such, the structure of perceiver effects might be different across levels of involvement and exposure.

With respect to *involvement*, people might have rather nuanced views about those who interact with them but simpler ones about people in general. For example, Peter might be a globally lenient judge but have high personal standards in terms of dependableness. Arguably, his issue with dependableness (a highly specific perceiver effect) would only come into play with people who surround him and not with people he knows from hearsay or from the news (lower specificity). To the contrary, it is conceivable that personality judgments at higher levels of involvement demand more cognitive effort and therefore pull for a simpler structure compared with judgments made at lower levels of involvement. For example, Peter might be highly concerned with the impression he makes in face-to-face contexts and might therefore hardly provide differentiated trait judgments about others (low specificity). In contrast, he might be more focused when providing judgments about people who are, for instance, presented on a screen and this might lead him to differentiate between traits more thoroughly (higher specificity).

With respect to *exposure*, it has been established that perceivers draw more heavily on general stereotypes when knowing little about targets and less so as exposure increases and as more individuating information becomes available (Biesanz, West, & Millevoi, 2007; Funder, 1999; Kenny, 2004). However, it is currently not clear what this means with respect to the structure of perceiver effects. On the one hand, the structure might become increasingly simple as exposure increases. For instance, Peter's starting point for judging others might be the expectation that people are mostly quite sociable and creative but often lack dependableness (high specificity) but after he gets a chance to learn more about targets' actual sociability, creativity, and dependableness, what remains is his tendency to be a globally lenient judge (low specificity). However, the opposite is also conceivable such that perceiver effects become increasingly nuanced with increasing exposure. Peter might see others as highly sociable, creative, and dependable at first glance (low specificity), but the more he learns about them the more might his high personal standards for dependableness play out and lead him to interpret others' behaviors as signs of low dependableness (high specificity).

Currently, little is known about contextual influences on the structure of perceiver effects, but perhaps they might account for the discrepant findings in the literature. Participants in Srivastava et al.'s (2010) studies were sitting together on a table when providing ratings and were thus personally involved but in Wood, Harms, et al.'s (2010) investigation (Study 2), questionnaires were completed anonymously at home and thus involvement was lower. Further, participants in Srivastava et al.'s (2010) studies were almost unacquainted with one another and thus exposure was lower than in Wood, Harms, et al.'s (2010) studies where participants were roommates. The present work addresses whether the context of judgments is a reason for the discrepancy in previous findings by exploring whether exposure and involvement affect the relative contributions of positivity and specificity in perceiver effects.

The Present Research

Perceiver effects are believed to be a ubiquitous component in person perception that have theoretical and practical importance. However, it is unclear whether perceiver effects follow a robust factor structure, how much they reflect a general positivity versus trait-specific tendencies, and whether findings vary systematically across trait frameworks and judgment contexts. To date, there is some empirical support for the positivity-specificity model in a Big Five framework, but this support is based on only two studies that took place in a single judgment context (Srivastava et al., 2010), and at least one other line of research using a different framework and different context partially contradicted those findings (Wood, Harms, et al., 2010). To better understand the general nature of perceiver effects, the present research analyzed personality judgments data from 10 studies (N = 2,199) that included different trait frameworks and judgment contexts.

The agenda of the present research was threefold. First, we tested competing a priori factor models in each dataset to examine the configural structure of perceiver effects (i.e., the necessity to model both global positivity and trait-specificity in addition to acquiescence bias). These analyses clarified whether perceiver effects reflect an undifferentiated tendency to see others as globally positive or negative or whether they also reflect tendencies to see others as high or low with respect to particular trait content. Second, we analyzed the factor solutions in more detail and quantified the contributions of acquiescence, global positivity, and trait-specificity to learn about the importance of these components relative to one another and relative to the overall differences occurring in personality judgments of a given trait. Third, to better understand the role of positivity and specificity, we also examined whether they influenced perceiver effects differently for different traits (the Big Five, agency & communion) and in different judgment contexts (level of exposure, level of involvement). This clarified whether perceiver effects have a constant psychological meaning or whether trait and contextual differences change the nature of perceiver effects.

Method

Overview of the Studies

We analyzed data from 10 studies, which are listed in Table 1, and all studies were in accordance with ethical guidelines of the respective institutions (Studies 1, 2, and 9: Office of Research Ethics of the University of Toronto [project title: Impression Study; protocol number: 00031168]; Study 3: German Research Foundation [project title: LE 2151/3–1; no protocol number]; Study 4: Office of Research Ethics of the University of Toronto [project title: Emerging Leaders Study; protocol number 00031108]; Study 5: McGill University Research Ethics Board [project title: Social Consequences of First Impressions; protocol number: 178–1015]; Study 6: Washington University in St. Louis IRB [project title: Blind Spots and Bright Spots; protocol number: 201105199]; Study 7: University of Münster FB7 Ethics Committee [project title: How exact is your perception of others′ personalities?; no protocol number], Study 8: Ethics guidelines at Humboldt-University of Berlin's Institute of Psychology did not require ethics approval for this nonexperimental, nonintrusive, and anonymous study; Study 10: Washington

Table 1
Study Details

Study number	Reference	*N* (perceivers)	*M* (*SD*) age	Average number of targets rated	% female	Design	Exposure to targets[a]	Involvement with targets[b]	Employed trait frameworks	Targets' task
1	Barranti & Carlson (2018a)	175	18.65 (.96)	4.0	77	H-B	low (1.1)	low (vid)	B5	Short unstructured getting acquainted conversation with another person (five minutes).
2	Barranti & Carlson (2018b)	137	19.66 (4.63)	4.0	78	H-B	low (1.1)	low (vid)	B5 & B2	Short unstructured getting acquainted conversation with another person (five minutes).
3	Leising, Locke, Kurzius, & Zimmermann (2016)	201	24.91 (5.05)	4.0	50	H-B	high (2.1)	low (vid)	B5 & B2	In the videos, targets went through a standardized ten-minute interview involving various tasks (reading aloud, humor, general knowledge, calculation skills, creativity, live events, etc.).
4	Carlson & Feinberg (2018)	123	19.78 (1.27)	4.5	54	R-R	low (1.1)	high (f2f)	B5 & B2	Within groups, each participant answered the question "if you had a superpower, what would it be and why?".
5	Heyman, Biesanz, & Human (2019)	543	20.42 (2.14)	5.4	85	R-R	low (1.1)	high (f2f)	B5	Short unstructured getting acquainted conversation in dyads (two-three minutes).
6	Wilson, Thompson, & Vazire (2017; Study 1)	124	20.19 (2.58)	3.5	67	R-R	low (1.1)	high (f2f)	B5	Short unstructured getting acquainted conversation in groups (five minutes).
7	Niemeyer, Back, & Nestler (2016)	146	22.72 (2.73)	4.7	77	R-R	high (2.1)	high (f2f)	B2	Participants first gave a quick self-introduction and then engaged in a leaderless group discussion about a moral dilemma.[c]
8	Rentzsch & Gebauer (2019)	421[d]	24.4 (6.71)	5.4	69	R-R	high (2.6)	high (f2f)	B5 & B2[e]	Participants worked together in groups over the course of three months and wrote a graded report at the end of the semester.
9	Barranti & Carlson (2017)	T1: 102 T2: 93 T3: 95	19.25 (1.91)	3.5	76	R-R	T1:low (.3) T2: low (.5) T3: high (1.7)	high (f2f)	B5	Participants played six rounds of a Public Goods Game on individual tablets. Mutual ratings were collected before the first round (T1), after the third round (T2), and after the sixth round (T3). After the fourth round, groups were invited to discuss their decisions but before that there were no conversations whatsoever.
10	Carlson (2016)	T1: 227 T2: 152	19.79 (.98)	4.3	62	R-R	T1: low (1.1) T2: high (2.6)	high (f2f)	B5	T1: Ice breaking game: Two Truths and a Lie T2: Three months later; after various class activities

Note. R-R = round-robin; H-B = half-block; B5 = Big Five framework; B2 = agency-communion framework.
[a] 0 = zero acquaintance (i.e., only physical appearance), 1 = very first impression (i.e., minimal single interaction), 2 = first impression (i.e., 2–4 chats), 3 = acquainted (i.e., 5 + chats). [b] vid = judgments based on videos, f2f = judgments based on face-to-face interactions. [c] Participants took part in two sessions in which they were part of a different group and perceiver effects from both sessions were averaged. [d] Round-robin ratings were assessed in two waves (around weeks 5 and 12 of participants' first semester), but not all participants completed both waves. We averaged perceiver effect scores of participants who completed both waves and used the single-wave scores of participants who completed only one wave. For a similar procedure with the same data see also Nehrlich, Gebauer, Sedikides, and Schoel (2019). [e] B2 judgments were provided by a reduced sample of $n = 294$.

University in St. Louis IRB [project title: Personality Perceptions; protocol number: 10–0832]). All studies recruited convenience samples with an academic background, and participants were either offered course credit or received monetary compensation. In all studies, perceivers were initially unacquainted with targets (i.e., perceivers who were already acquainted with targets were excluded), with the exception of Study 8 which was a two-wave classroom study that took place around week 5 and 12 of the participants' first semester and thus involved a certain level of prior acquaintance. Additional materials including data, R-Code, supplementary analyses, and a list of prior publications using data presented here are retrievable from osf.io/kr5ms/.[2] None of the prior publications has addressed the structure of perceiver effects.

Designs. Studies 1–3 used a *half-block design* where each perceiver judged the same set of four targets (i.e., perceivers and targets were distinct), and Studies 4–10 used a *round-robin design* where participants in small groups rated and were rated by group members (i.e., perceivers were also targets; see the Design column in Table 1). Studies 9 and 10 were longitudinal round-robin studies where perceivers provided ratings at several stages of the getting-acquainted process. Note that the distinction between half-block and round-robin designs is crucial for how perceiver effects are computed (see section Operationalizing Perceiver Effects below).

Involvement. In Studies 1–3 perceivers watched targets on videos, whereas in Studies 4–10 participants were involved with targets face-to-face. Video and face-to-face studies were classified as low and high involvement studies, respectively (see the Involvement column in Table 1).

Exposure. In all studies, perceivers rated targets who engaged in some kind of social interaction. The nature of these interactions differed across studies and time points and are outlined in the Target's Task column in Table 1. Thus, perceivers made judgments on the basis of quite different information across studies and time points, which was rated on a scale from 0 to 3 (0 = *physical appearance only*, 1 = *minimal single interaction*, 2 = *two to four chats*, 3 = *five or more chats*) by 10 independent raters. Interrater agreement was excellent (ICC[3, 10] = .96) and ratings were averaged across raters. Exposure ratings were classified as *low* and *high*, when they were below and above the midpoint of the scale, respectively (see the Exposure column in Table 1). Notably, participants in the first stage of Study 10 played the icebreaking game "Two Truths and a Lie," which is the same game that was played in Srivastava et al.'s Study 1 (2010).

Trait frameworks. In five of the 10 studies we analyzed personality judgments exclusively within a Big Five framework, and in one of the studies we analyzed judgments exclusively within a Big Two framework. In the remaining four studies we adopted both frameworks (see the Employed Trait Framework column in Table 1).

Measures

In all studies, perceivers provided ratings about targets on several personality characteristics some of which came from established personality scales. Specifically, the BFI-10 (Rammstedt & John, 2007) was used in Study 8, the Big Five measure from Borkenau and Ostendorf (1998) was used in Study 3, the Interpersonal Adjective List (Jacobs & Scholl, 2005) was used in Studies 3 and 7, and the TIPI (Gosling et al., 2003) was used in Studies 6 and 10. Moreover, many studies assessed additional items, and they varied widely with respect to the content and number of these items. From this heterogenous pool of items, we needed to make a selection with which we could analyze the structure of perceiver effects. Given that such selection affords many researcher degrees of freedom and can undermine the replicability of the obtained results, we preregistered our item selections prior to processing and analyzing the data (osf.io/2fkw5/). In the following, we will describe how we selected items at this stage and explain why the preregistered item sets turned out to be unworkable in some instances. We will then describe how we handled these instances by further narrowing down the considered items in a second step.

In the first selection step, the first and second author screened all of the available items and selected the ones that appeared to tap into one of the Big Five or Big Two factors while discarding those that were unrelated or ambiguously related to the Big Five or Big Two. We avoided ambiguous items because cross-loadings would likely limit our ability to disentangle positivity variance and trait-specific variance in perceiver effects.[3] For example, we did not include the item "this person is a leader" because leadership perceptions might reasonably reflect perceptions of extraversion (e.g., good leadership owing to high enthusiasm), conscientiousness (good leadership owing to high dependability), emotional stability (good leadership owing to high resilience), as well as agency (good leadership owing to high assertiveness) and communion (good leadership owing to high commitment). At the same time, leadership perceptions might of course reflect perceived global positivity (good leadership due to a generally favorable personality). As such, we selected items that seemed to index single trait factors and discarded items that implied cross-loadings (e.g., "this person is a leader") and items that were irrelevant to the Big Five or Two trait content (e.g., "I respect this person"). Items were screened separately for the Big Five and the Big Two frameworks and then preregistered. For each item, we computed perceiver effect scores (cf. section Operationalizing Perceiver Effects), which we then factor analyzed (cf. section Modeling).

Although the preregistered item sets fit the data well for the Big Two framework, many of the Big Five models did not converge despite several respecifications. These problems tended to occur in studies with large and imbalanced sets of items (e.g., one study had 14 indicators for one trait factor but only four indicators for another trait factor) but not in studies with smaller and more balanced sets of items (e.g., some studies had two indicators for each trait factor). To test whether item selection was the problem rather than perceiver effect structure, we modeled self-reports for the same item sets and found that self-report models converged even worse than the perceiver effect models (for a detailed docu-

[2] All of the studies reported here were not originally designed to investigate the structure of perceiver effects and featured many aspects unrelated to the present research (e.g., self-report scales, behavioral tasks, performance measures, etc.). For clarity, shared research materials on the OSF exclusively pertain to the interpersonal perception portion of the studies. Shared data include aggregated person-level data but no raw dyad-level data for privacy reasons.

[3] It would have also been possible to analyze all the items while using a more exploratory approach that freely estimates all possible cross-loadings. Whereas such an approach has the preferable feature of exploiting all the information available in the data, it hampers the interpretability of the substantive factors and would have made it difficult to disentangle positivity and trait-specificity.

mentation of all initial analyses refer to OS1). These results suggest that our original Big Five item selections were too heterogeneous and imbalanced.[4]

In light of these insights, we deviated from the preregistration and went through a second round of item selection. Specifically, we reduced the number of items to 10 (i.e., two indicators per Big Five factor) in each study, and aimed to obtain measures as similar as possible to established Big Five short scales like the TIPI (Gosling et al., 2003) or the BFI-10 (Rammstedt & John, 2007) to increase comparability among our studies as well as with published studies that use Big Five short scales. To reduce the Big Five item sets in an unbiased way, the first author conducted a mini survey among five personality psychology experts who were unfamiliar with the preliminary results. For orientation, the experts were provided with the original items of the TIPI (Gosling et al., 2003) and asked to indicate which of the available items would yield the most suitable 10 item Big Five measure. Consensus was quite high in most cases (e.g., *sympathetic* was suggested as an indicator of agreeableness by all experts) but when recommendations were ambivalent we computed parcels of the respective items (e.g., two experts named *punctual* and three experts named *keeps promises* as their preferred indicator of conscientiousness). The final item sets are displayed in Appendix A, and a detailed documentation of the item reduction process including the results of the expert survey can be found in the additional materials (OS2).

Reducing the number of items had a disadvantage. Namely, each latent trait was modeled with two indicators only. At the same time, however, reducing the number of items also had two key advantages—it successfully resolved the problems of convergence and it also increased the comparability of the results across data sets. We ran a series of control analyses to gauge the weight of the disadvantage. Specifically, there were two studies in which larger sets of established Big Five items had been administered (30 items in Study 3, 15 items in Study 5) and which allowed us to estimate well-identified and balanced factor models in addition to the 10-item models. The results were virtually identical across the considered item sets, suggesting that the use of only two indicators per latent trait did not systematically bias the results of the 10-item models. For brevity, we present details about the control analyses including data and R-code in the additional materials (OS3) and focus on the 10-item models in the main part of the article.

Operationalizing Perceiver Effects

According to the Social Relations Model (SRM; Kenny, 1994), any interpersonal judgment decomposes into the perceiver effect (how the perceiver judges others in general), the target effect (how the target is judged by others in general), and the relationship effect (dyad-specific perceptions including measurement error). Before modeling the factor structure of perceiver effects, three central questions need to be answered: First, how influential are perceiver effects? Do they contribute substantially to the overall variance in interpersonal perceptions at all or are perceptions only driven by target and relationship differences? Second, how can the perceiver effect for a given perceiver on a given item be scored? Third, what is the reliability of this score? Below, we will address these questions separately for half-block and round-robin designs.

Perceiver effects in half-block data. For half-block data we ran two-way random effects models with random intercepts for perceivers and targets (Judd, Westfall, & Kenny, 2017) to estimate the relative contribution of perceiver effects to the total variation in interpersonal judgments. In line with the assumptions of the SRM, this model (also referred to as the crossed-random effects model) treats perceivers and targets as independent random sources of variation. We ran the model with the *lme4* package (Version 1.1–21; Bates, Mächler, Bolker, & Walker, 2015) in *R* (R Development Core Team, 2008) for each item and divided the intercept variance for perceivers by the sum of the intercept variances for perceivers and targets and the residual variance. Thus, this coefficient can take values ranging from 0 to 1 and reflects the percentage of variance in *single* judgments that is attributable to differences between perceivers. Following SRM terminology, we will refer to this coefficient as *perceiver variance*. To score perceiver effects on each item, we computed means across all judgments coming from the same perceiver across targets (Kenny, 1994). Note that, because every perceiver judged the same set of targets, target variance does not contribute to variation in these scores (for a mathematical proof, see additional material OS4). Thus, the resulting scores are comprised of reliable perceiver differences and of random error and the proportion of reliable variance can be indexed by Cronbach's coefficient alpha. We will refer to this as *perceiver effect reliability*.

Perceiver effects in round-robin data. In round-robin data it is not possible to rely on simple crossed-random effects models, average scores, and coefficient alpha to estimate perceiver variances, perceiver effect scores, and their reliabilities. The reason for this is that every perceiver rates a slightly different set of targets which leads to complex dependencies and a risk of confounding perceiver and target influences. We therefore ran social relations analyses (for an accessible overview see Back & Kenny, 2010) as implemented in the *R* (R Development Core Team, 2008) package *TripleR* (Version 1.5.3; Schönbrodt, Back, & Schmukle, 2012) to estimate these properties for each item. We retained the default of the package which removes any group with fewer than four members[5] and report standardized perceiver variance estimates which have the same interpretation as the perceiver variance coefficients calculated from a crossed-random effects model described above. Further, we saved the estimated perceiver effect scores, which are free from systematic target influences and are thus comparable to the average scores from half-block studies. We also report the corresponding reliability estimates provided by TripleR but we note that these are not strictly comparable to coefficient alpha because they additionally account for the number of round-robin groups (Bonito & Kenny, 2010).

[4] Besides the reduction of items in the Big Five analyses, other deviations from the preregistration became necessary over the course of the project. For example, one dataset that was mentioned in the preregistration turned out to be unsuitable because it used a rating instruction that undermined perceiver variance and three of the datasets reported here (Studies 2, 5, and 8) were not mentioned in the preregistration because they became available after the preregistration was completed. Overall, it seems more appropriate to characterize the present work as not being preregistered—at least not in a strict sense.

[5] This procedure may reduce the effective sample size. The sample sizes we report here might therefore not be perfectly identical to the sample sizes reported in other publications that analyzed the same data but used different software packages or different settings.

Table 2
Perceiver Variances (Perceiver Effect Reliabilities) by Trait and Study

Study	Big Five framework						Big Two framework		
	A	C	E	ES	O	Study average	Agy	Com	Study average
1	.11 (.49)	.16 (.47)	.06 (.38)	.13 (.41)	.24 (.63)	.14 (.47)	—	—	—
2	.18 (.54)	.22 (.53)	.08 (.33)	.21 (.52)	.31 (.69)	.20 (.52)	.19 (.55)	.21 (.58)	.20 (.57)
3	.11 (.44)	.07 (.31)	.13 (.45)	.09 (.38)	.12 (.49)	.10 (.41)	.10 (.41)	.15 (.52)	.13 (.46)
4	.47 (.80)	.46 (.80)	.36 (.76)	.50 (.81)	.52 (.85)	.46 (.80)	.42 (.79)	.67 (.90)	.55 (.85)
5	.32 (.72)	.27 (.68)	.09 (.45)	.16 (.51)	.27 (.69)	.22 (.61)	—	—	—
6	.42 (.75)	.45 (.76)	.06 (.33)	.48 (.79)	.32 (.66)	.34 (.66)	—	—	—
7	—	—	—	—	—	—	.15 (.55)	.17 (.59)	.16 (.57)
8	.15 (.52)	.17 (.60)	.09 (.45)	.14 (.51)	.16 (.54)	.14 (.52)	.21 (.64)	.22 (.61)	.21 (.62)
9 (T1)	.55 (.82)	.4 (.72)	.35 (.67)	.56 (.80)	.54 (.80)	.48 (.76)	—	—	—
9 (T2)	.62 (.84)	.62 (.86)	.40 (.72)	.58 (.82)	.63 (.85)	.57 (.82)	—	—	—
9 (T3)	.50 (.79)	.54 (.81)	.41 (.76)	.60 (.83)	.47 (.73)	.50 (.78)	—	—	—
10 (T1)	.33 (.67)	.36 (.75)	.11 (.51)	.38 (.75)	.33 (.72)	.30 (.68)	—	—	—
10 (T2)	.32 (.72)	.34 (.74)	.14 (.56)	.35 (.70)	.27 (.67)	.28 (.68)	—	—	—
Trait average	.34 (.67)	.34 (.67)	.19 (.53)	.35 (.65)	.35 (.69)	.31 (.64)	.21 (.59)	.28 (.64)	.25 (.61)

Note. Perceiver variances reflect the percentage of variance in single judgments that are attributable to differences between perceivers (rather than targets or dyads). Perceiver effect reliabilities reflect the percentage of reliable variance in aggregated perceiver effect scores. The presented estimates were averaged across items relating to the same trait. A = agreeableness; C = conscientiousness; E = extraversion; ES = emotional stability; O = openness; Agy = agency; Com = communion.

Perceiver variances and perceiver effect reliabilities. Table 2 displays the perceiver variances and perceiver effect reliabilities for each of the 10 studies, averaged across items of Big Five and Big Two traits, respectively. Across studies and traits, perceiver effects accounted for roughly 30% of the overall variance in person impressions. For extraversion and agency, perceiver variance was somewhat descriptively lower than for the remaining traits. Further, perceiver variance was descriptively lower in video studies compared with face-face-studies, especially to those involving low exposure (e.g., Studies 4 and 9). Importantly, perceiver effects largely accounted for at least 10% of the overall variance which is commonly regarded as a benchmark for a substantial contribution (Kenny, 1994). Differences across studies and traits were also mirrored in the corresponding perceiver effect reliabilities. Because the reliabilities for perceiver effect scores also account for the aggregation across targets, they were necessarily larger than the perceiver variances of the respective study and trait. However, they were overall modest in size given that scores were based on single-item ratings about a handful of targets and they cannot be subjected to the reliability standards of questionnaire scales. Item level statistics are provided in the additional materials (OS5).

Modeling

We fitted a series of CFA models to the perceiver effect data, and a detailed protocol of the fitting procedure (OS6) as well as the data and R-scripts required to reproduce the analyses (OS7) are provided in the additional materials. As described in the introduction, we considered three potential models to account for covariation of perceiver effects across items: A unidimensional positivity-only model, a trait-specificity model with correlated trait factors, and a higher-order positivity-specificity model. Statistically, the unidimensional model is nested within the higher-order model, which again is nested within the correlated factors model. Conceptually, this is reflected by different degrees of parsimony versus flexibility. The positivity-only model is highly parsimonious (i.e., inflexible) by claiming that there is a single perceiver characteristic, how positively others are seen generally, that can describe all the covariation in perceiver effects of different items. In contrast, the trait-specificity model is highly flexible (i.e., nonparsimonious) by claiming trait-specific perceiver effects and remaining silent about the relations between them. Finally, the positivity-specificity model offers a compromise between parsimony and flexibility by both assuming some trait-specificity but also offering an explanation for covariation among trait factors, namely a common positivity factor.[6]

Within the Big Two framework, we compared the positivity-only model and a two-factor model with correlated factors, which equally reflects the trait-specificity and the positivity-specificity model because there are only two trait factors. Technically, a higher-order factor is not identified with a two-factor model, and the only parameter to be estimated is the covariation between the two factors. Assuming that perceptions of agency and communion correlate to the degree they share the feature of being evaluative, we consider a two-factor solution as evidence for the positivity-specificity model and interpret the factor correlation in terms of individual differences in global positivity versus negativity.

Finally, paralleling previous research (Srivastava et al., 2010), we included a method factor to account for differences in acquiescent responding. This is accomplished by specifying a factor that

[6] There are alternative specifications that are conceptually highly similar to the higher-order model. For instance, one could adjust the unidimensional model by allowing correlated residuals among indicators relating to the same trait dimensions or one might specify orthogonal trait factors and an additional orthogonal positivity factor (i.e., a bifactor model). In the present case, where there are just two indicators per trait, all of these models are equivalent with respect to fit to the data. Therefore, these alternative factor models are not further considered here. Interested readers are referred to Gignac (2016) and Mansolf and Reise (2017) for a discussion of the relation between the bifactor and the higher-order model.

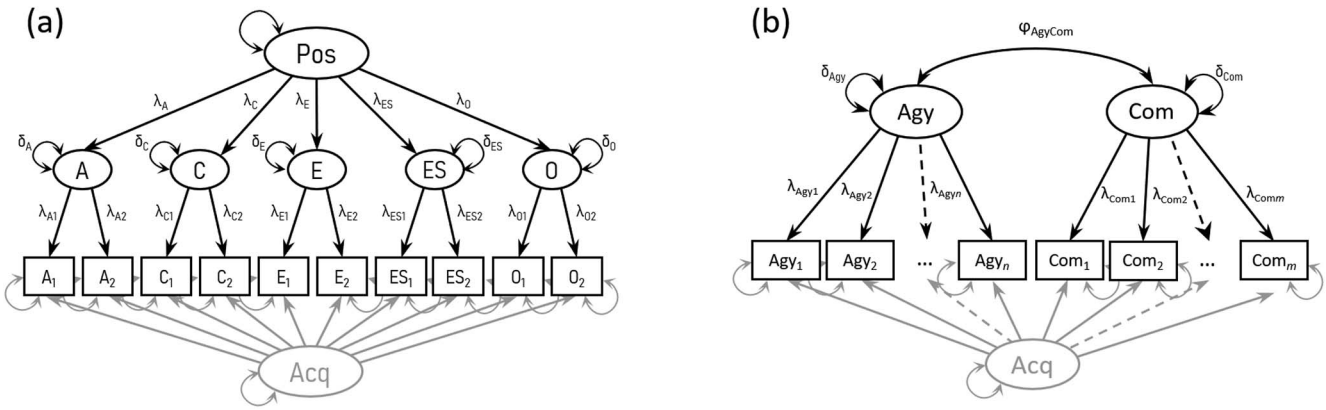

Figure 2. Fully specified positivity-specificity model including a method factor for acquiescent responding (all loadings fixed to 1) for the Big Five framework (a) and for the Big Two framework (b).

has a fixed association of 1 with all indicators (including both regularly and reverse keyed indicators) and that is uncorrelated with all other factors in the model (Billiet & McClendon, 2000). An acquiescence factor removes variance in perceiver effects that gets introduced when some respondents more than others endorse all items irrespective of their content. In the Big Two analyses of Studies 3, 5, and 8, all items were keyed in the same direction with respect to valence and thus, an acquiescence factor could not be modeled. In the remaining models, the additional model complexity introduced by the acquiescence factor (one variance parameter) was justified by respective increases in model-fit. Figure 2 displays the positivity-specificity model for the Big Five and the Big Two framework featuring an acquiescence factor.

We used the *R* (R Development Core Team, 2008) package *lavaan* (Version 0.6–2; Rosseel, 2012) with full information maximum likelihood estimation. All residuals were constrained to be uncorrelated, and cross-loadings were not allowed with one exception—two items in Study 7, which had different item stems (one relating to a person's gestures and the other one to their facial expressions) but an identical item body (i.e., confident-dominant). We report all parameter estimates from a fully standardized solution where both the observed and latent variables' variances equal 1. Given that the sample sizes vary across studies, we refrained from comparing models based on likelihood ratio significance tests, because this would introduce bias favoring parsimonious models in studies with lower power and bias favoring flexible models in studies with higher power. Instead, we based model comparisons on the Bayesian information criterion (BIC; Schwarz, 1978), which is insensitive to differences in sample sizes and controls for the number of free parameters in a model (i.e., it involves a penalty for nonparsimony).

Of note, by comparing the three theoretically derived models, our analytical strategy focused on model selection rather than model fit. Nevertheless, we report the comparative fit index (CFI), the root-mean-square error of approximation (RMSEA), and standardized root-mean-square residual (SRMR) in addition to the BIC to allow for a comprehensive evaluation of model fit. Given that the use of strict cutoff values for good fit (e.g., CFI $>$ .950; RMSEA $<$ .080; SRMR $<$ .080) is disputable on various grounds (West, Taylor, & Wu, 2012), we also present the fit indices from Srivastava et al. (2010) and from self-report models of the same item sets as additional benchmarks.

Handling of Improper Solutions

In some instances, the estimator initially yielded improper solutions. These included negative residual variances of observed variables, negative variances of trait factors in the positivity-specificity model, and factor correlations above 1 in the trait-specificity model. All of the negative residual variances of observed variables were not significantly different from zero suggesting that improper solutions were a result of random sampling fluctuations (Kolenikov & Bollen, 2012) and accordingly, we fixed the respective parameters to zero to obtain a normal solution. With respect to negative trait-specific variances in the positivity-specificity model, some of the estimates were significantly different from zero ($p <$ .05) suggesting that the model might be overfactored (Rindskopf, 1984). To avoid overfactoring and to achieve a normal solution, we fixed the respective trait-specific variances to zero which effectively leaves the positivity factor to explain all of the covariation between the respective items and which is statistically equivalent to removing the respective trait factor from the model. Appendix B shows which parameters were fixed to zero. It is important to note that these types of modifications are conceptually uncontroversial: Observed variables with an error variance of zero can be interpreted as being perfectly explained by the factors in the model, and trait factors with a specific variance of zero can be interpreted as resembling nothing but differences in positivity.

However in the third case, when correlations between trait factors were estimated above 1 in the trait-specificity model, there is in fact a conceptual problem. Here, the only way to avoid overfactoring and to obtain a normal solution would have been to combine trait factors. However, it is questionable whether and how such combined factors can be substantively interpreted because no conceptual guidance exists for the resulting Big Four or Big Three models. Therefore, we do not report fit statistics for those models.

Details on the fitting procedure (e.g., the exact order in which parameters were fixed and models were refitted) are documented for all models and all item sets in the additional materials (OS6).

Results

Is There a Robust Configural Structure of Perceiver Effects?

Overall, model comparisons suggested that perceiver effects reflect both positivity and trait-specificity in most contexts; that is, results favored the Positivity-Specificity Model. Table 3 displays the fit indices of the different theoretical models sorted by trait framework and exposure. In all models, the signs of factor loadings were in the expected direction. The BIC favored the positivity-specificity model in 14 of 17 cases and the positivity-only model in the remaining three cases (see the bold-printed indices in Table 3). These three cases all concerned Big Five judgments that were based either on physical appearance or on a minimal single interaction (low exposure) in face-to-face interactions (high involvement). The trait-specificity model yielded improper solutions in five of 12 cases and even when a normal solution was found, the BIC consistently favored the more parsimonious positivity-specificity model.[7] Detailed model solutions including factor loadings and trait specific residual variances of the positivity-specificity model in each study are presented in Appendix B.

In terms of model fit, the overall pattern of fit statistics might be considered borderline as some models passed conventional criteria of good model fit whereas others failed to do so. However, fit indices for the positivity-specificity model were comparable to the ones reported by Srivastava et al. (2010), $\chi^2(df) = 110\ (29)$; $CFI = .85$; $RMSEA = .094$; $SRMR = .069$. To ensure that borderline model fit was not due to perceiver effect structure, we also fitted a five factor model to self-reports on the same Big Five items and a two factor model to self-reports on the same Big Two items to compare indices of absolute fit. Similar to the perceiver effect models, self-report models were borderline, $.815 \leq \text{CFIs} \leq .947$, $.059 \leq \text{RMSEAs} \leq .117$, $.052 \leq \text{SRMRs} \leq .098$ (see OS6 in the additional materials for results in each study). In light of these benchmarks, the perceiver effect models appear to fit the data reasonably well.

How Strongly Do Positivity and Specificity Contribute to Perceiver Effects?

Given that the model comparisons suggested that perceiver effects reflect both positivity and trait-specificity in most contexts, we next sought to better understand how strongly these components influenced perceptions of others. To directly index the relative importance of global positivity and trait-specificity, we performed variance decomposition for each item in each study.[8] Equation 1a displays the variance decomposition for an item y_{it}, where λ_{it} is the item's loading on the associated lower-order trait factor, λ_t is the respective trait factor's loading on the higher order positivity factor, var(pos) is the higher-order factor's variance, $\text{var}(\delta_t)$ is the lower order factor's residual or trait-specific variance, λ_{acq} is the item's loading on the acquiescence factor, var(acq) is the variance of the acquiescence factor, and $\text{var}(\varepsilon_{it})$ is the residual variance. In a standardized solution, where $\text{var}(y_{it})$, var(pos), and var(acq) equal 1, the formula reduces to Equation 1b. For Big Two models the decomposition is the same except that λ_t is replaced by ϕ_{AgyCom} and that $\text{var}(\delta_t)$ is not a residual variance.

$$\text{var}(y_{it}) = \lambda_{it}^2 \cdot \lambda_t^2 \cdot \text{var(pos)} + \lambda_{it}^2 \cdot \text{var}(\delta_t) + \lambda_{acq}^2 \cdot \text{var(acq)} + \text{var}(\varepsilon_{it}) \quad (1a)$$

$$1 = \lambda_{it}^2 \cdot \lambda_t^2 + \lambda_{it}^2 \cdot \text{var}(\delta_t) + \lambda_{acq}^2 + \text{var}(\varepsilon_{it}) \quad (1b)$$

In the standardized case, the first additive term $\lambda_{it}^2 \cdot \lambda_t^2$ provides a direct estimate of the percentage of item variation driven by differences in positivity bias, the second term $\lambda_{it}^2 \cdot \text{var}(\delta_t)$ provides a direct estimate of the percentage of item variation driven by trait specific perceiver tendencies, and the third additive term λ_{acq}^2 provides a direct estimate of item variation driven by acquiescence bias.[9] These variance proportions index the percentage of *perceiver variance* that is attributable to positivity, specificity and acquiescence, respectively. An alternative way of scaling these proportions is in reference to the item's *overall variance*. This is achieved by multiplying the variance proportions based on Equation 1b with the perceiver variance estimate. These proportions then refer to the component's total influence on a rating by a single rater about a single target. Thus, they reveal how influential the perceiver effect component of interest (e.g., positivity bias) is in relation to *all* other judgment influences including target and relationship characteristics, other perceiver effect components (e.g., trait specificity and acquiescence), and measurement error. The results of variance decomposition for the Big Five and the Big Two framework are presented in Figure 3.

The results show that perceiver effects carry a substantial portion of signal above and beyond acquiescence and measurement error. That is, they are a reflection both of the tendency to evaluate others positively versus negatively and of the tendency to see others in particular ways for specific traits. Together, these two components accounted for 43% and 45% of perceiver variance in judgments of the Big Five and the Big Two, respectively (Figure 3a). When taking into account how much perceiver variance items elicited, this corresponded to 12% and 10% of the overall differences in personality judgments (Figure 3b). It is noteworthy that the ratio of positivity and specificity

[7] We also explored the possibility of a five-plus-two meta-trait model in the 12 datasets that featured Big Five judgments. In six cases, the estimator did not converge or yielded improper solutions that could not be fixed by reasonable respecifications which lines up with reports by Srivastava et al. (2010). In the remaining cases, three meta-trait models were inferior to the respective positivity-specificity model and three were superior as indicated by higher and lower BICs, respectively. In the meta-trait models that were preferable, the two higher-order factors were highly correlated (*r*s between .55 and .71), suggesting noticeable perceiver differences in global positivity. In light of these results, we did not deem the five-plus-two meta-trait model a robust and viable alternative to the positivity-specificity model.

[8] The three datasets where the positivity-only model was preferable were retained in these analyses. Although the unidimensional configural structure of these data implies that trait-specific factors explain little variance, it does not imply a particular amount of positivity variance (e.g., relative to error and acquiescence) and thus, there is still a need for quantification.

[9] Because acquiescence could not be modeled in Big Two analyses of Studies 2, 4, and 8, we used the estimates from the studies' respective Big Five analyses (where negatively worded items had been present) instead. This builds on the fact that acquiescence is by definition a response tendency that affects every item in the same way regardless of content and avoids artificial inflation of positivity variance.

Table 3
Model-Fit of the Positivity-Only, Trait-Specificity, and Positivity-Specificity Model Sorted by Trait Framework and Level of Exposure to Targets

					Positivity-only model			Trait-specificity model			Positivity-specificity model		
Framework	Exposure[a]	Involvement[b]	Study	No. of items	$\chi^2(df)$	CFI; RMSEA; SRMR	BIC	$\chi^2(df)$	CFI; RMSEA; SRMR	BIC	$\chi^2(df)$	CFI; RMSEA; SRMR	BIC
B5	0	hi	9 T1	10	84 (34)	.859; .121; .086	**3025**	—			80 (32)	.865; .122; .090	3030
B5	1	hi	9 T2	10	81 (34)	.878; .123; .067	**2788**	—			71 (30)	.894; .122; .064	2796
B5	1	hi	4	10	75 (34)	.837; .099; .086	**4969**	55 (24)	.877; .103; .072	4997	62 (30)	.872; .094; .077	4975
B5	1	hi	5	10	559 (34)	.743; .169; .090	8593	174 (24)	.926; .108; .056	8272	199 (29)	.917; .104; .061	**8265**
B5	1	hi	10 T1[c]	10	155 (34)	.745; .126; .086	8736	—			93 (30)	.867; .097; .070	**8696**
B5	1	lo	1	10	141 (34)	.831; .134; .090	5022	70 (24)	.926; .106; .102	5003	90 (29)	.904; .110; .087	**4997**
B5	1	lo	2	10	122 (34)	.827; .138; .088	4211	50 (24)	.948; .090; .054	4188	66 (29)	.927; .097; .068	**4179**
B5	2	hi	6	10	134 (34)	.693; .154; .107	4742	—			70 (31)	.879; .101; .081	**4693**
B5	2	lo	3	10	254 (34)	.452; .180; .124	3168	62 (24)	.904; .090; .078	3029	97 (32)	.838; .101; .095	**3021**
B5	2	hi	9 T3	10	63 (34)	.919; .096; .085	2903	—			45 (31)	.961; .070; .085	**2898**
B5	3	hi	10 T2	10	220 (34)	.644; .190; .101	5639	65 (25)	.923; .103; .072	5529	94 (32)	.881; .113; .084	**5523**
B5	3	hi	8	10	332 (34)	.684; .145; .081	4170	94 (24)	.926; .083; .047	3992	116 (29)	.908; .085; .055	**3984**
B2	1	hi	4	6	23 (9)	.936; .115; .049	3016				9 (8)	.993; .040; .033	**3007**
B2	1	lo	2	9	178 (27)	.826; .203; .065	3350				53 (26)	.968; .088; .039	**3230**
B2	2	lo	3	8	291 (19)	.454; .267; .166	2532				87 (18)	.861; .139; .102	**2333**
B2	2	hi	7	16	561 (102)	.747; .176; .114	2642				302 (101)	.889; .117; .084	**2388**
B2	3	hi	8	9	426 (27)	.807; .224; .120	3648				285 (26)	.875; .184; .087	**3513**

Note. For each line, the lowest BIC is displayed in bold print to indicate which of the three models is the preferred one. B5 = Big Five framework; B2 = Agency and Communion framework. [a] 0 = zero acquaintance (i.e., only physical appearance), 1 = very first impression (i.e., minimal single interaction), 2 = first impression (i.e., 2–4 chats), 3 = acquainted (i.e., 5 + chats). [b] lo = indirect assessment based on videos, hi = direct assessment based on face-to-face interactions —: Improper solution. [c] Close replication of Srivastava, Guglielmo, and Beer (2010).

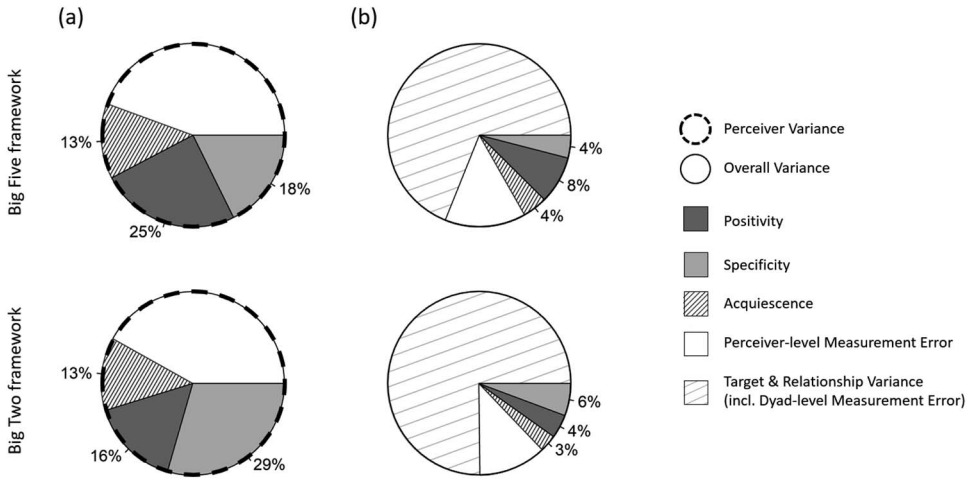

Figure 3. Relative contributions of global positivity, trait-specificity, and acquiescence in relation to perceiver variance (a), and in relation to the overall variance in personality judgments (b) of the Big Five and the Big Two.

is not necessarily equal when computed in reference to perceiver variance versus in reference to overall variance, because the computations are made separately for each item and are then aggregated. For the Big Five, items that yielded a lot of positivity variance in perceiver effects also yielded more perceiver variance relative to target and relationship variance, which explains why positivity was more predominant over specificity in the left as compared with the right hand side of the top panel of Figure 3.

Does the Relative Importance of Positivity and Specificity Vary Across Trait Domain and Judgment Context?

Also in Figure 3, the relative contributions of positivity and specificity were different for the two trait frameworks. Whereas positivity was predominant in perceiver effects of Big Five judgments, specificity was predominant in perceiver effects of Big Two judgments. This raises the question of whether perceiver effects might reflect different judgment tendencies depending on the considered trait content.

Positivity versus specificity: A matter of which trait is being judged? We computed variance proportions of global positivity and trait-specificity separately for the traits within the Big Five and the Big Two framework. To provide a focal illustration of the relative importance of positivity and specificity, we rescaled the two proportions so they would sum up to 100% and plotted them in pie charts that do not feature acquiescence and error variance. The results are displayed in Figure 4.

There were considerable differences in the relative contributions of positivity and specificity across the Big Five traits. For extra-

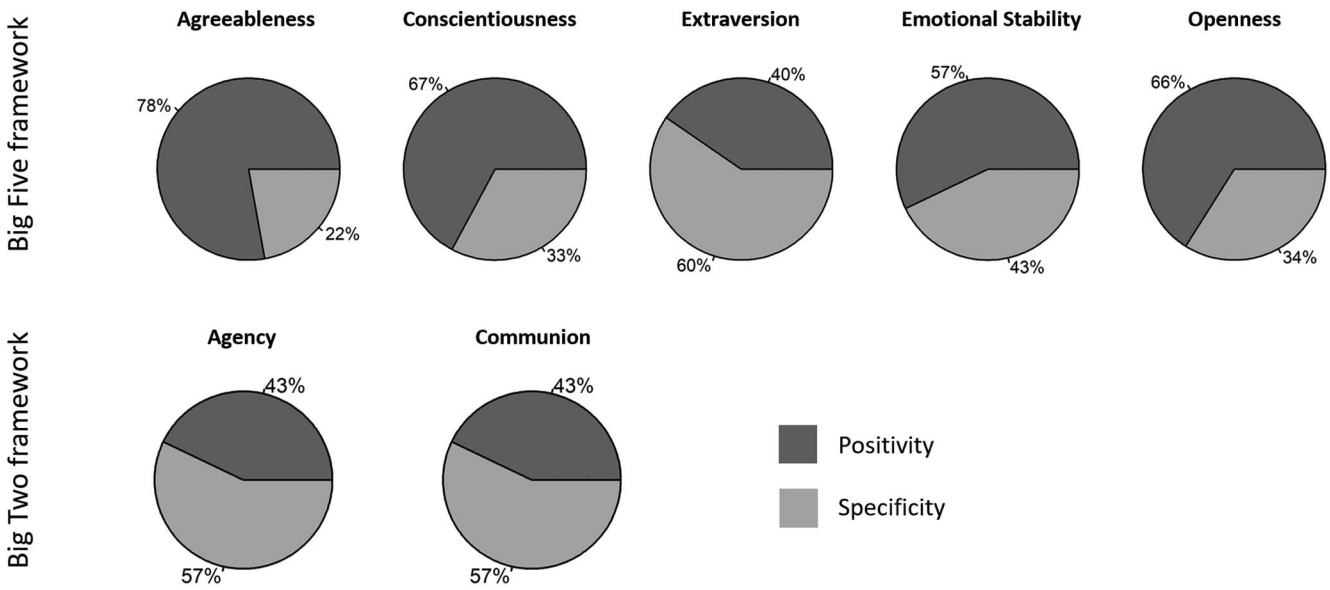

Figure 4. Ratio of global positivity and trait-specificity in perceiver effects of the Big Five and the Big Two.

version, perceiver effects were largely driven by trait-specific tendencies but for agreeableness they were largely driven by global positivity and the other traits fell in between. Within the Big Two framework, no such differences could be observed. The reason for this is that, in a two factor model, positivity is estimated from the factor correlation, which means that the ratio between positivity and specificity is necessarily equal for both factors. Hence, the agency-communion framework cannot meaningfully speak to trait-level differences in the relative importance of positivity and specificity.

Positivity versus specificity: A matter of judgment contexts? To examine whether contextual features render perceiver effects to be more or less nuanced, we aggregated variance proportions of positivity and specificity separately for studies which featured low versus high involvement with targets and for studies which featured low versus high exposure to targets (cf. Table 1). Figure 5 displays the results.

The proportion of specificity was descriptively much larger in studies which featured low involvement with targets (i.e., video studies) compared with studies which featured high involvement with targets (i.e., face-to-face studies) in both trait frameworks. Concerning exposure, the proportion of trait-specificity was descriptively larger in studies which featured higher exposure to targets compared with studies which featured lower exposure to targets. This difference was modest in the Big Five framework but considerable in the Big Two framework.

Discussion

In the present research we set out to examine the structure of perceiver effects using a large and heterogeneous pool of personality judgments data featuring 10 studies with a total of $N = 2{,}199$ participants. Overall, findings strongly suggest that, above and beyond response bias, perceiver effects reflect two psychologically meaningful constructs: the tendency to globally evaluate others positively (vs. negatively) and the tendency to see others as high (vs. low) with respect to particular trait content. A model taking both of these components into account outperformed a simpler positivity-only and a less parsimonious trait-specificity model in the vast majority of studies. The positivity-only model was preferable only in Big Five judgments assessed after brief personal encounters. The findings also include a replication of Srivastava et al.' Study 1 (2010) because one of our studies (Study 10 [T1]) closely mirrored the original study (i.e., same trait measure, same context) and found the same results. In a broader sense, the findings strongly demonstrated the positivity-specificity model's generalizability to various trait frameworks and judgment contexts.

How Much Do Positivity and Specificity Matter?

Beyond identifying the configural structure of perceiver effects, one of our goals was to learn about the relative importance of response bias, global positivity, and trait-specificity. Extending prior research, we decomposed the variance of perceiver effects using the factor solutions of the positivity-specificity model. Overall, this revealed that positivity and specificity are both similarly important and that their influence is larger than the influence of acquiescence bias. Thus, even without a tendency to generally use high numbers in their ratings, people might have high perceiver effects on a given rating dimension (e.g., an average rating of 9 on a scale from 1 to 10 on the dimension of sociability) because they have a positive evaluative attitude toward others generally (e.g., because they like most people) and/or because they see others as especially sociable, outgoing, and so forth (e.g., they have a personal issue with extraversion).

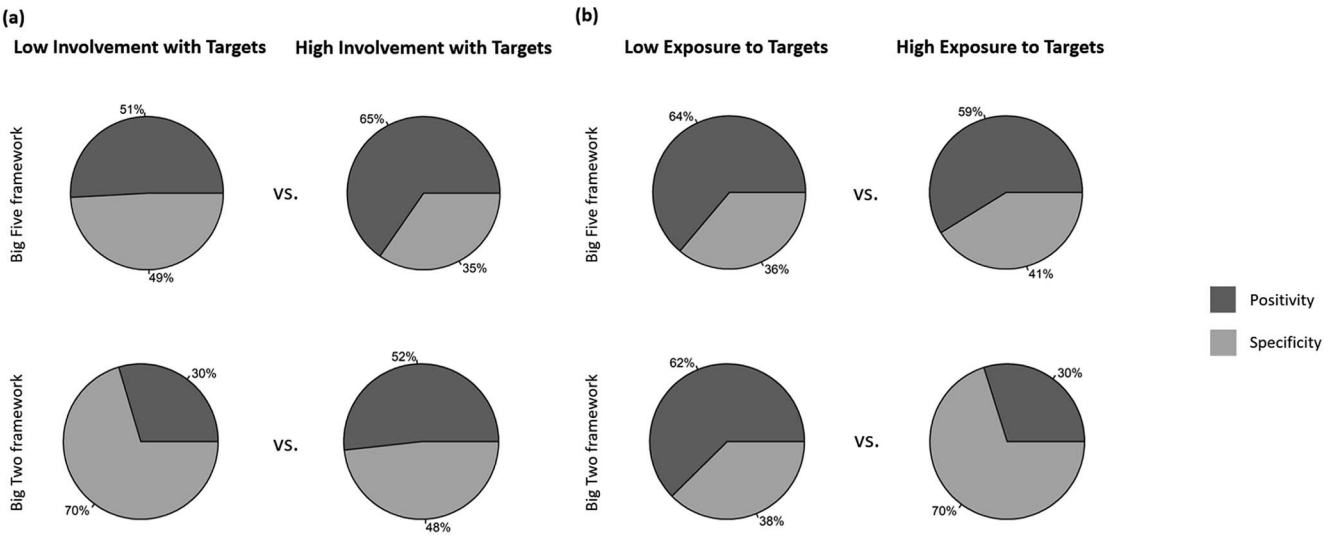

Figure 5. Ratio of global positivity and trait-specificity in perceiver effects by level of involvement with targets (a) and level of exposure to targets (b). Low involvement studies featured judgments based on videos, and high involvement studies featured judgments based on face-to-face interactions. Low exposure studies featured judgments based on physical appearance or based on a minimal single interaction and high exposure studies featured judgments based on two to four chats or more.

When Do Positivity and Specificity Matter? The Influence of Trait Domain and Judgment Context

Arguably, what is more important than the general contributions of positivity and specificity are their contributions with respect to particular trait content and in particular judgment contexts. We conceived that positivity and specificity might affect perceiver effects differently across trait domains and across contexts that involve different levels of involvement and exposure. Indeed, we found such differences.

Trait differences. The relative contributions of positivity and specificity were quite different across the Big Five traits, and two aspects of this finding are particularly interesting. First, positivity was particularly dominant for agreeableness. Agreeableness has been identified as the most *evaluative* trait within the Big Five taxonomy (Funder & Dobroth, 1987; John & Robins, 1993), and most aspects of agreeableness are closely tied to likability (Koch, Alves, Krüger, & Unkelbach, 2016; Suitner & Maass, 2008; Wortman & Wood, 2011). Thus, an evaluative attitude toward others in general is probably what underlies other-perceptions of agreeableness and this clarifies the conceptual meaning of the positivity factor. Individuals with high positivity may assume that others are agreeable (and to a lesser degree also open, emotionally stable, etc.) simply because they *like* most others (cf. Leising, Scherbaum, Locke, & Zimmermann, 2015).

The finding that more evaluative traits are more strongly influenced by positivity can also account for discrepancies in previous studies. Srivastava et al. (2010) used a measure that explicitly avoids evaluatively extreme items (Gosling et al., 2003) and found evidence of trait-specificity, whereas Wood, Harms, et al. (2010) used a measure with many evaluative items (Wood, Nye, et al., 2010) and found no evidence of trait-specificity. Hence, the fact that the latter investigation was geared more toward capturing evaluative attitudes may explain the absence of trait-specificity in this specific study setup.

The second interesting aspect concerns the fact that specificity was particularly dominant in perceiver effects of extraversion. Extraversion has been identified as the most *observable* Big Five trait which typically causes extraversion judgments by unacquainted observers to reach relatively high consensus (i.e., target variance; Funder & Dobroth, 1987; John & Robins, 1993). A typical consequence of this is a reduction of perceiver variance, which was indeed observable in the present data (see Table 2). Taken together, this shows that higher observability reduces positivity but does not reduce specificity. As an illustration, this means that Peter's perception of others' extraversion is largely independent of his overall tendency to like or dislike others but is still subjected to his idiosyncratic way of interpreting observed behaviors as signs of intro- or extraversion.

Differences in evaluativeness and observability might also explain why we found less positivity and more specificity in the Big Two compared with the Big Five framework. Agency, much like extraversion, is highly observable and communion, much like agreeableness, is highly evaluative. Correspondingly, we found less perceiver variance in impressions of agency than in impression of communion (see Table 2). If the high observability rendered perceiver effects for agency rather trait-specific, less global positivity could possibly be found in a Big Two framework given that positivity was estimated by the overlap between agency and communion and not by a higher-order factor. For instance, if perceiver effects of agency were perfectly trait-specific, the correlation with the communion factor would be zero. This would have been interpreted as absence of positivity although in reality, the perceiver effects for communion could largely be driven by positivity. As a solution, one could conceptualize agency as a dimension that carries evaluatively relevant content but communion as evaluation per se, a suggestion that has recently been made by Imhoff and Koch (2017). It will be a challenge for future research to further qualify the psychological meaning of the positivity factor and explore the (im)possibility of dissociating it from other-perceptions of high warmth, cooperativeness, and morality, that is, from defining features of agreeableness and communion. For example, future work could examine how strongly a higher-order positivity factor in perceiver effects is associated with evaluative attitudes toward targets (i.e., liking, physical attraction), state affect (i.e., mood), or prosocial decision making (i.e., trusting).

Context differences. Finally, we explored whether the relative importance of positivity and specificity depended on the level of perceivers' personal involvement with targets and on the level of exposure to targets. We found that specificity variance was stronger when personal involvement was low, suggesting that perceiver effects might reflect rather nuanced views about others when targets are judged from a distance but that these views become simpler once perceivers are personally involved with targets. Perhaps, perceivers' cognitive capacities are lower in face-to-face interactions than when they are personally disengaged and, in turn, perceivers might rely more heavily on their overall gut feeling about others. However, these observations were only descriptive and were based on a limited number of studies. Future work that employs an experimental manipulation of cognitive load may offer a more formal test of this idea and provide more fine-grained insights into underlying mechanisms.

We also observed that specificity variance was higher at higher levels of exposure to targets. This suggests that the personal issues perceivers have with respect to particular trait content come more and more into play as perceivers acquire knowledge about targets. A possible explanation for this is that these issues indeed reflect idiosyncratic *interpretations* of observed behaviors instead of mere *expectations* and that a certain level of exposure is required for such interpretations to play out. An experiment that manipulates the amount and ambiguity of information about targets could clarify the role of interpretative processes.

Implications

The present research has important theoretical and practical implications. Theoretically, results provide support for classic theories proposing that individuals have highly generalized working models of others (Bowlby, 1988; Erikson, 1959, 1968) and for theories that conceptualize models of others to operate on the level of traits as well (e.g., assumed similarity or projection; Campbell et al., 1964; Cronbach, 1955). Given that our results suggest that idiosyncratic perceptions of others are equally important both on a continuum from negative to positive and on the level of trait-content, it will be important to better understand how both aspects of perceiver effects develop and how they change for individuals over time. For example, positivity biases may develop as part of

the overall attachment system whereas trait-specific perceiver effects may result from content-specific social projection effects.

The positivity-specificity model of perceiver effects might also inform theories of basic personality or personality pathology. For example, in the domain of normal personality, people's beliefs about what others are like might explain their behavior. Peter might be generous and sociable with others because he generally sees the best in others, but he might avoid relying on others because he assumes others are low in dependableness. With respect to clinical theories, one basic assumption is that people higher in personality pathology often have overly negative views of others, but a more complete theory of personality pathology might involve understanding both overly negative perceptions of others across trait domains as well as trait-specific perceptions. For example, depressive pathology might be characterized by a global negativity bias (Beck, 1979) but obsessive–compulsive, pathology might pertain more specifically to skepticism about the qualities of others' performances (American Psychiatric Association, 2013). Further, the maladaptive interpersonal signatures by borderline, narcissistic, passive-aggressive, and psychopathic personalities have been theorized to be rooted in biased other-perceptions in terms of agency and communion (Hopwood, 2018).

Practically, our results suggest that researchers and practitioners should model both global and trait-specific perceiver variance. Conceptualizing and measuring perceiver effects as a unidimensional construct will often result in misleading or misguided conclusions. As such, researchers who want to examine people's idiosyncratic views about others in a particular trait domain should control for perceiver effects in other domains to avoid positivity variance to confound their conclusions. On the other hand, researchers who want to measure people's tendencies to see others positively versus negatively should average across impressions of several traits and the more evaluative these traits are, the fewer of them they will need to obtain a reliable measure. Overall, disentangling the unique sources of perceiver variance will be an important step for measuring personality correlates of perceiver effects (e.g., do people with different personalities also have different blueprint of others' personalities?) as well as identifying interpersonal consequences of perceiver effects (e.g., which blueprints affect the development and maintenance of social relationships?).

Limitations and Future Directions

One caveat of the present work is that the positivity-specificity model showed only moderate fit to the data in absolute terms. Although our primary focus was on model selection, the indication of misfit might still be cause for concern and might suggest minor misspecifications. Therefore, model solutions may not be particularly robust and the corresponding variance proportions should rather be regarded as rough estimates than being taken at face value. Attempts to model perceiver effects in larger sets of items and in larger samples in the future may increase the robustness and precision of the analyses pioneered here.

More generally, the main goal of the present work was to understand how simple versus nuanced people's idiosyncratic perceptions of others are, but we did not identify the sources of these idiosyncrasies. At least two plausible sources are attachment styles and self-perceptions. Concerning the attachment system, it is conceivable that individuals who are securely attached perceive others more positively in general, but to our knowledge no research has directly addressed this hypothesis to date. Concerning self-views, the idea that people might see others as they see themselves has been extensively studied under the label of *assumed similarity* (Cronbach, 1955) but, despite implying a trait-specific phenomenon, assumed similarity correlations have habitually been calculated without controlling for global positivity in other-perceptions (for a recent review see Thielmann, Hilbig, & Zettler, 2018). Forming a noticeable exception, Srivastava et al. (2010) found that assumed similarity existed even when positivity was controlled for, suggesting that people might in fact project specific aspects of their own personalities onto others. However, this evidence is still preliminary and more research is needed to identify potential roots of trait-general and trait-specific perceiver effects.

Another potential caveat of the present work is that the psychological meaning of perceiver effects can be affected by how much perceivers have the chance to interact with targets independently of other perceivers. Indeed, in some studies, we were unable to disentangle the degree to which perceiver effects reflected judgment tendencies versus actual reactions from others; that is, a social reality (cf. Rau, Nestler, Dufner, & Nestler, 2019). For example, dominant individuals tend to elicit more submissive behavior from others (Markey, Funder, & Ozer, 2003; Sadler & Woody, 2003) and thus, a dominant participant might have seen others as submissive not because he or she has an idiosyncratic perception but because others were in fact especially submissive around him or her. In principle, perceivers have to interact with targets one-on-one for this to apply because then (a) the perceiver has an opportunity to influence the target, and (b) the target's behavior is not observed by the other perceivers. Studies 5, 8, and 10 [T2] featured such interactions and thus it is possible that perceiver effects captured not only idiosyncratic perceptions but also a social reality. In contrast, perceivers could not influence targets' behavior in video studies (Studies 1–3) and they all had access to targets' behavior in group interaction studies (Studies 4, 6, 7, 9, and 10 [T1]). Although results did not systematically differ between these sets of studies, future research might still examine the degree to which interacting with targets independently of other perceivers influences perceiver effects. For example, it might be possible to videotape and later code behavior in dyadic interactions to test if perceiver effects indeed reflect actual behavior.

Finally, we found a robust factor structure of perceiver effects across various setups but there are additional ways to test generalizability. First, studying further judgment domains (e.g., motives, intentions, goals) and studying contexts that feature higher levels of involvement (e.g., dating) and higher levels of exposure (e.g., judgments of close friends) would be valuable extensions. Second, future research should examine to which degree people bring their perceiver effects with them across contexts using rotation designs (i.e., each perceiver is part of several distinct groups) or ambulatory assessment techniques (i.e., repeated measures in daily situations). For example, such approaches might reveal that people who tend to see others as sociable in cooperative situations tend to do so in competitive situations as well or that those who globally assume the best about others when they first meet new people continue to do so once they are getting closer with them. This methodology might also allow for an explicit, experimental test of the role of judgment contexts, which we were unable to provide given the low number of studies within a given context. Finally, future work should examine

perceiver effects in different populations such as among middle aged or older adults and in different cultures.

Conclusion

If Peter perceives a specific person as highly sociable, this can be attributable to actual qualities of this person, to the relationship he has with this person, or to Peter himself. The present research demonstrated that the latter component, Peter's perceiver effect, again compounds two psychologically meaningful constructs: a tendency to view most other people positively or negatively and a tendency to view them as more or less sociable in particular. In some contexts, it will largely reflect the first tendency, positivity, and in other contexts it will largely reflect the second tendency, specificity. Thus, to learn what it is that Peter brings into his judgment, it is crucial to consider what trait is being judged and to what degree Peter is involved with and informed about the target person.

References

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93,* 751–763. http://dx.doi.org/10.1037/0022-3514.93.5.751

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.

Back, M. D., & Kenny, D. A. (2010). The social relations model: How to understand dyadic processes. *Social and Personality Psychology Compass, 4,* 855–870. http://dx.doi.org/10.1111/j.1751-9004.2010.00303.x

Bakan, D. (1966). *The duality of human existence. An essay on psychology and religion*. Chicago, IL: Rand McNally.

Barranti, M., & Carlson, E. N. (2017, June). *Predicting moral behaviors: Does the self or do others know best?* Poster Presented at the Biannual Meeting of the Association for Research in Personality, Sacramento, CA.

Barranti, M., & Carlson, E. N. (2018a). *Pilot Project for the UTM Impression Study*. Unpublished Manuscript.

Barranti, M., & Carlson, E. N. (2018b). *The UTM Impression Study*. Unpublished Manuscript.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. Advance online publication. http://dx.doi.org/10.18637/jss.v067.i01

Beck, A. T. (1979). *Cognitive therapy of depression*. New York, NY: Guilford Press.

Biesanz, J. C., West, S. G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintance and consensus and self-other agreement in judgments of personality. *Journal of Personality and Social Psychology, 92,* 119–135. http://dx.doi.org/10.1037/0022-3514.92.1.119

Billiet, J. B., & McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling, 7,* 608–628. http://dx.doi.org/10.1207/S15328007SEM0704_5

Bonito, J. A., & Kenny, D. A. (2010). The measurement of reliability of social relations components from round-robin designs. *Personal Relationships, 17,* 235–251. http://dx.doi.org/10.1111/j.1475-6811.2010.01274.x

Borkenau, P., & Ostendorf, F. (1998). The Big Five as states: How useful is the Five-Factor Model to describe intraindividual variations over time? *Journal of Research in Personality, 32,* 202–221. http://dx.doi.org/10.1006/jrpe.1997.2206

Bowlby, J. (1988). *A secure base*. New York, NY: Basic Books.

Campbell, D. T., Miller, N., Lubetsky, J., & O'Connell, E. J. (1964). Varieties of projection in trait attribution. *Psychological Monographs: General and Applied, 78,* 1–33. http://dx.doi.org/10.1037/h0093872

Carlson, E. N. (2016). Meta-accuracy and relationship quality: Weighing the costs and benefits of knowing what people really think about you. *Journal of Personality and Social Psychology, 111,* 250–264. http://dx.doi.org/10.1037/pspp0000107

Carlson, E. N., & Feinberg, M. (2018). *Interpersonal perceptions in an Introduction to Organizational Behavior class*. Unpublished Dataset.

Cronbach, L. J. (1955). Processes affecting scores on understanding of others and assumed similarity. *Psychological Bulletin, 52,* 177–193. http://dx.doi.org/10.1037/h0044919

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. http://dx.doi.org/10.1037/0022-3514.91.6.1138

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73,* 1246–1256. http://dx.doi.org/10.1037/0022-3514.73.6.1246

Dufner, M., Leising, D., & Gebauer, J. E. (2016). Which basic rules underlie social judgments? Agency follows a zero-sum principle and communion follows a non-zero-sum principle. *Personality and Social Psychology Bulletin, 42,* 677–687. http://dx.doi.org/10.1177/0146167216640902

Erikson, E. H. (1959). *Identity and the life cycle*. New York, NY: International Universities Press.

Erikson, E. H. (1968). *Identity, youth and crisis*. New York, NY: Norton.

Fraley, C. R. (2002). Attachment stability from infancy to adulthood: Meta-analysis and dynamic modeling of developmental mechanisms. *Personality and Social Psychology Review, 6,* 123–151. http://dx.doi.org/10.1207/S15327957PSPR0602_03

Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception*. San Diego, CA: Academic Press.

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52,* 409–418. http://dx.doi.org/10.1037/0022-3514.52.2.409

Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence, 55,* 57–68. http://dx.doi.org/10.1016/j.intell.2016.01.006

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. http://dx.doi.org/10.1016/S0092-6566(03)00046-1

Hehman, E., Sutherland, C. A. M., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. *Journal of Personality and Social Psychology, 113,* 513–529. http://dx.doi.org/10.1037/pspa0000090

Heyman, J. L., Biesanz, J. C., & Human, L. J. (2019). *Online and offline social experiences: Associations between social media use and in-person first impressions*. Manuscript in preparation.

Hopwood, C. J. (2018). Interpersonal dynamics in personality and personality disorders. *European Journal of Personality, 32,* 499–524. http://dx.doi.org/10.1002/per.2155

Hopwood, C. J., Schade, N., Krueger, R. F., Wright, A. G. C., & Markon, K. E. (2013). Connecting *DSM–5* personality traits and pathological beliefs: Toward a unifying model. *Journal of Psychopathology and Behavioral Assessment, 35,* 162–172. http://dx.doi.org/10.1007/s10862-012-9332-3

Imhoff, R., & Koch, A. (2017). How orthogonal are the Big Two of social perception? On the curvilinear relation between agency and communion. *Perspectives on Psychological Science, 12,* 122–137. http://dx.doi.org/10.1177/1745691616657334

Jacobs, I., & Scholl, W. (2005). Interpersonale Adjektivliste (IAL). *Diagnostica, 51,* 145–155. http://dx.doi.org/10.1026/0012-1924.51.3.145

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluative-

ness, and the unique perspective of the self. *Journal of Personality, 61,* 521–551. http://dx.doi.org/10.1111/j.1467-6494.1993.tb00781.x

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology, 68,* 601–625. http://dx.doi.org/10.1146/annurev-psych-122414-033702

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

Kenny, D. A. (2004). PERSON: A general model of interpersonal perception. *Personality and Social Psychology Review, 8,* 265–280. http://dx.doi.org/10.1207/s15327957pspr0803_3

Koch, A., Alves, H., Krüger, T., & Unkelbach, C. (2016). A general valence asymmetry in similarity: Good is more alike than bad. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42,* 1171–1192. http://dx.doi.org/10.1037/xlm0000243

Kolenikov, S., & Bollen, K. A. (2012). Testing negative error variances. *Sociological Methods & Research, 41,* 124–167. http://dx.doi.org/10.1177/0049124112442138

Leary, T. (1957). *Interpersonal diagnosis of personality*. New York, NY: Ronald Press.

Leising, D., Locke, K. D., Kurzius, E., & Zimmermann, J. (2016). Quantifying the association of self-enhancement bias with self-ratings of personality and life satisfaction. *Assessment, 23,* 588–602. http://dx.doi.org/10.1177/1073191115590852

Leising, D., Scherbaum, S., Locke, K. D., & Zimmermann, J. (2015). A model of "substance" and "evaluation" in person judgments. *Journal of Research in Personality, 57,* 61–71. http://dx.doi.org/10.1016/j.jrp.2015.04.002

Mansolf, M., & Reise, S. P. (2017). When and why the second-order and bifactor models are distinguishable. *Intelligence, 61,* 120–129. http://dx.doi.org/10.1016/j.intell.2017.01.012

Markey, P. M., Funder, D. C., & Ozer, D. J. (2003). Complementarity of interpersonal behaviors in dyadic interactions. *Personality and Social Psychology Bulletin, 29,* 1082–1090. http://dx.doi.org/10.1177/0146167203253474

Nehrlich, A. D., Gebauer, J. E., Sedikides, C., & Schoel, C. (2019). Agentic narcissism, communal narcissism, and prosociality. *Journal of Personality and Social Psychology, 117,* 142–165. http://dx.doi.org/10.1037/pspp0000190

Niemeyer, L. M., Back, M. D., & Nestler, S. (2016). *EXACT*. Retrieved from https://osf.io/78h9n/

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41,* 203–212. http://dx.doi.org/10.1016/j.jrp.2006.02.001

Rau, R., Nestler, S., Geukes, K., Back, M. D., & Dufner, M. (2019). Can other-derogation be beneficial? Seeing others as low in agency can lead to an agentic reputation in newly formed face-to-face groups. *Journal of Personality and Social Psychology, 117,* 201–227. http://dx.doi.org/10.1037/pspp0000250

Rau, R., Nestler, W., Dufner, M., & Nestler, S. (2019). *Seeing the best or worst in others: A measure of generalized other-perceptions*. Manuscript under review.

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rentzsch, K., & Gebauer, J. E. (2019). On the popularity of agentic and communal narcissists: The tit-for-tat hypothesis. *Personality and Social Psychology Bulletin, 45,* 1365–1377. http://dx.doi.org/10.1177/0146167218824359

Rindskopf, D. (1984). Structural equation models: Empirical identification, Heywood cases, and related problems. *Sociological Methods & Research, 13,* 109–119. http://dx.doi.org/10.1177/0049124184013001004

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48,* 1–36. http://www.jstatsoft.org/v48/i02/

Sadler, P., & Woody, E. (2003). Is who you are who you're talking to? Interpersonal style and complementarity in mixed-sex interactions. *Journal of Personality and Social Psychology, 84,* 80–96. http://dx.doi.org/10.1037/0022-3514.84.1.80

Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality, 69,* 847–879. http://dx.doi.org/10.1111/1467-6494.696167

Saucier, G., Ostendorf, F., & Peabody, D. (2001). The non-evaluative circumplex of personality adjectives. *Journal of Personality, 69,* 537–582. http://dx.doi.org/10.1111/1467-6494.694155

Schönbrodt, F. D., Back, M. D., & Schmukle, S. C. (2012). TripleR: An R package for social relations analyses based on round-robin designs. *Behavior Research Methods, 44,* 455–470. http://dx.doi.org/10.3758/s13428-011-0150-4

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464. http://dx.doi.org/10.1214/aos/1176344136

Srivastava, S., Guglielmo, S., & Beer, J. S. (2010). Perceiving others' personalities: Examining the dimensionality, assumed similarity to the self, and stability of perceiver effects. *Journal of Personality and Social Psychology, 98,* 520–534. http://dx.doi.org/10.1037/a0017057

Suitner, C., & Maass, A. (2008). The role of valence in the perception of agency and communion. *European Journal of Social Psychology, 38,* 1073–1082. http://dx.doi.org/10.1002/ejsp.525

Thielmann, I., Hilbig, B. E., & Zettler, I. (2018). Seeing me, seeing you: Testing competing accounts of assumed similarity in personality judgments. *Journal of Personality and Social Psychology*. Advance online publication. http://dx.doi.org/10.1037/pspp0000222

Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology, 98,* 281–300. http://dx.doi.org/10.1037/a0017908

West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). New York, NY: Guilford Press.

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37,* 395–412. http://dx.doi.org/10.1037/0022-3514.37.3.395

Wilson, R. E., Thompson, R. J., & Vazire, S. (2017). Are fluctuations in personality states more than fluctuations in affect? *Journal of Research in Personality, 69,* 110–123. http://dx.doi.org/10.1016/j.jrp.2016.06.006

Wood, D., Harms, P., & Vazire, S. (2010). Perceiver effects as projective tests: What your perceptions of others say about you. *Journal of Personality and Social Psychology, 99,* 174–190. http://dx.doi.org/10.1037/a0019390

Wood, D., Nye, C. D., & Saucier, G. (2010). Identification and measurement of a more comprehensive set of person-descriptive trait markers from the English lexicon. *Journal of Research in Personality, 44,* 258–272. http://dx.doi.org/10.1016/j.jrp.2010.02.003

Wortman, J., & Wood, D. (2011). The personality traits of liked people. *Journal of Research in Personality, 45,* 519–528. http://dx.doi.org/10.1016/j.jrp.2011.06.006

Appendix A

Final Item Selection

Study	Big Five					Big Two	
	A	C	E	ES	O	Agy	Com
1	1: sympathetic 2: compassionate[a]	1: self-disciplined 2: disorganized (r)	1: outgoing 2: happy	1: calm 2: irritable[a] (r)	1: intellectual 2: creative	—	—
2	1: warm 2: critical (r)	1: self-disciplined 2: disorganized (r)	1: introverted (r) 2: timid (r)	1: calm 2: anxious (r)	1: creative 2: intelligent	1: dominant 2: influence 3: is a leader 4: high status	1: cooperative 2: warm 3: compassionate 4: fair 5: trustworthy
3	1[P]: good-natured/ considerate 2: dogmatic (r)	1: responsible 2: reckless (r)	1: sociable 2: reserved (r)	1[P]: emotionally stable/ relaxed 2[P]: touchy/vulnerable (r)	1: witty 2: fanciless (r)	1: self-assured 2: assertive 3: silent (r) 4: bashful (r)	1: cordial 2: empathetic 3: hostile (r) 4: malicious (r)
4	1: warm 2: selfish (r)	1: hardworking 2: disorganized (r)	1: happy 2: introverted (r)	1: calm 2: anxious (r)	1: intelligent 2: creative	1: has influence 2: is a leader 3: high status	1: warm 2: gives praise 3: trustworthy
5	1: helpful and unselfish 2: starts quarrels (r)	1[P]: follows through/ is a reliable worker 2: can be careless (r)	1: outgoing, sociable 2: reserved (r)	1: relaxed 2[P]: calm/tense (r)	1: original 2: a deep thinker	—	—
6	1: sympathetic, warm 2: critical, quarrelsome (r)	1: dependable, self-disciplined 2: disorganized, careless (r)	1: extraverted, enthusiastic 2: reserved, quiet (r)	1: calm, emotionally stable 2: anxious, easily upset (r)	1: open, complex 2: conventional, uncreative (r)	—	—
7	—	—	—	—	—	1: assertive 2: shy (r) 3: confident expression 4: plain statements 5: hesitant statements (r) 6: loud voice 7: confident movements 8: tense expression (r) 9: shaky voice (r)	1: empathic 2: hostile (r) 3: contemptuous expression (r) 4: cooperative statements 5: approachable movements 6: friendly voice 7: affirmative expression
8	1: trusting 2: finds fault with others (r)	1: does a thorough job 2: lazy (r)	1: outgoing, sociable 2: reserved (r)	1: relaxed 2: nervous (r)	1: active imagination 2: few artistic interests (r)	1: ambitious 2: commanding 3: dominant 4: is a leader	1: warm 2: compassionate 3: honest 4: caring 5: sympathetic
9	1: kind 2: critical (r)	1[P]: punctual/keeps promises 2: self-disciplined	1: outgoing 2: lethargic (r)	1: calm 2: worried (r)	1: intellectual 2: creative	—	—
10	1: sympathetic, warm 2: critical, quarrelsome (r)	1: dependable, self-disciplined 2: disorganized, careless (r)	1: extraverted, enthusiastic 2: reserved, quiet (r)	1: calm, emotionally stable 2: anxious, easily upset (r)	1: open, complex 2: conventional, uncreative (r)	—	—

Note. All items were rated on Likert-type scales. For brevity, some items are displayed in a shortened but semantically synonymous variant. Study 3 used Big Five items from Borkenau and Ostendorf (1998) and Big Two items from the Interpersonal Adjective List (Jacobs & Scholl, 2005). Studies 6 and 10 used the TIPI (Gosling, Rentfrow, & Swann, 2003). Study 8 used the BFI-10 (Rammstedt & John, 2007). Items used in the remaining studies were each composed from various instruments by the original authors. (r) = reverse coded.

[a] In Study 1, *critical* was suggested by the experts as a (negative) indicator of agreeableness but showed unexpected correlations with all other items and prevented the estimator from converging. Further *worried* was suggested as a (negative) indicator of Emotional Stability but was too highly correlated with *happy*, an indicator of Extraversion, and prevented the estimator from converging. As a solution, *compassionate* and *irritable* were post hoc selected as replacements, respectively.

[P] Parceled because of ambivalence in experts' suggestions.

(*Appendices continue*)

Appendix B

Table B1. Standardized Parameter Estimates From the Positivity-Specificity Model Within a Big Five Framework

	Trait factor loadings										Positivity factor loadings					Trait specific residual variances				
Study	λ_{A1}	λ_{A2}	λ_{C1}	λ_{C2}	λ_{E1}	λ_{E2}	λ_{ES1}	λ_{ES2}	λ_{O1}	λ_{O2}	λ_{A}	λ_{C}	λ_{E}	λ_{ES}	λ_{O}	δ_{A}	δ_{C}	δ_{E}	δ_{ES}	δ_{O}
1	.60	.62	.65	−.59	.46	.73	.30	−.67	.72	.27	.61	.74	.68	.93	.81	.62*	.45*	.53*	.14	.34
2	.58	−.31	.67	−.74	−.79	−.63	.53	−.73	.70	.57	.95	.67	.72	.81	.80	.09	.55*	.48*	.35*	.37*
3	.75	−.66	.97	−.38	.98	−.48	.97	−.50	.70	−.78	.47	.52	.45	.32	.86	.78*	.73*	.79*	.90*	.25
4[a]	.54	−.70	.57	−.59	.59	−.30	.42	−.62	.78	.42	.90	1	.88	.63	.67	.18	0	.23	.60*	.56*
5	.80	−.72	.83	−.63	.85	−.42	.81	.79	.83	.60	.96	.96	.43	.51	.78	.07	.08	.81*	.74*	.40*
6	.57	−.61	.63	−.70	.67	−.86	.81	−.72	.87	−.49	1	.74	.42	1	.49	0	.46*	.82*	.0	.76*
8	.74	−.33	.78	−.69	.78	−.61	.57	−.89	.78	−.61	.78	.68	.74	.59	.74	.39*	.54*	.45*	.65*	.46*
9 T1[a]	.58	.16	.45	.47	.37	−.52	.57	−.43	.74	.25	1	1	.26	.93	1	0	0	.93*	.14	0
9 T2[a]	.53	.30	.76	.69	.69	−.56	.74	−.48	.79	.64	1	.98	.82	.78	.91	0	.05	.33	.39*	.17
9 T3	.66	−.21	.47	.53	.58	−.81	.65	−.72	.52	.41	.69	1	.80	.78	1	.53	0	.35*	.39*	0
10 T1	.77	−.47	.72	−.59	.63	−.67	.65	−.72	.50	−.65	1	.67	.49	.83	.81	0	.54*	.76*	.32*	.34*
10 T2	.64	−.63	.94	−.53	.93	−.60	.69	−.86	.81	−.59	1	.69	.59	.68	.75	0	.53*	.65*	.54*	.44*

Note. Trait specific residual variances of 0 were fixed to avoid improper solutions. In these cases, the positivity factor loading is necessarily 1.
[a] Model comparison favored the positivity-only model (cf. Table 1).
* $p < .05$.

Table B2. Standardized Parameter Estimates From the Positivity-Specificity Model Within the Big Two Framework

	Trait factor loadings															Factor correlation
Study	λ_{A1}	λ_{A2}	λ_{A3}	λ_{A4}	λ_{A5}	λ_{A6}	λ_{A7}	λ_{A8}	λ_{A9}	λ_{C1}	λ_{C2}	λ_{C3}	λ_{C4}	λ_{C5}	λ_{C6}	λ_{C7}	φ_{AC}
2	.30	.95	.89	.84						.85	.74	.90	.85	.72			.77*
3	.59	.48	−.78	−.78						.45	.51	−.77	−.80				−.08
4	.77	.88	.17							.60	.83	.67					.81*
7	.66	−.59	.61	.71	−.70	.78	.62	−.69	−.68	.56	−.59	−.55	.72	.78	.86	.83	.64*
8	.73	.86	.44	.53						.90	.94	.78	.88	.90			.69*

* $p < .05$.

Appendix C

Overview of Additional Materials Retrievable From osf.io/kr5ms/

OS1: Fitting Protocol for Initial Item Sets
OS2: Expert Survey for Item Selection
OS3: Control Analyses for Studies 3 and 5
OS4: Validity Check for Mean Scores as Estimates of Perceiver Effects in Half-Block Designs
OS5: Item-Level Perceiver Variances and Perceiver Effect Reliabilities
OS6: Fitting Protocol for Final Item Sets
OS7: Data and R-Code
OS8: List of Prior Publications Using Data Presented in the Article
OS9: Original Materials

Received March 24, 2019
Revision received November 11, 2019
Accepted November 15, 2019 ■