Personality and Social Psychology Bulletin

Personality Judgments Based on Physical Appearance
Laura P. Naumann, Simine Vazire, Peter J. Rentfrow and Samuel D. Gosling

The online version of this article can be found at:

Published by:

⑤SAGE

http://www.sagepublications.com

On behalf of:

**SP
SP**

Society for Personality and Social Psychology, Inc.

Additional services and information for *Personality and Social Psychology Bulletin* can be found at:

Email Alerts: http://psp.sagepub.com/cgi/alerts

Subscriptions: http://psp.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations http://psp.sagepub.com/cgi/content/refs/35/12/1661

Personality Judgments Based on Physical Appearance

Laura P. Naumann
University of California, Berkeley

Simine Vazire
Washington University in St. Louis

Peter J. Rentfrow
University of Cambridge

Samuel D. Gosling
University of Texas at Austin

Despite the crucial role of physical appearance in forming first impressions, little research has examined the accuracy of personality impressions based on appearance alone. This study examined the accuracy of observers' impressions on 10 personality traits based on full-body photographs using criterion measures based on self and peer reports. When targets' posture and expression were constrained (standardized condition), observers' judgments were accurate for extraversion, self-esteem, and religiosity. When targets were photographed with a spontaneous pose and facial expression (spontaneous condition), observers' judgments were accurate for almost all of the traits examined. Lens model analyses demonstrated that both static cues (e.g., clothing style) and dynamic cues (e.g., facial expression, posture) offered valuable personality-relevant information. These results suggest that personality is manifested through both static and expressive channels of appearance, and observers use this information to form accurate judgments for a variety of traits.

Keywords: accuracy; person perception; personality; appearance; facial expression

We Wheelwrights do not scoff at the appearance of things. Things often *are* as they appear. First impressions matter.

A Prayer for Owen Meany, John Irving, 1989

Judgments based on physical appearance are ubiquitous and consequential. In face-to-face interactions, appearance is the first piece of information available to others and it can powerfully influence perceivers' subsequent behavior (Efran, 1974; Snyder, Tanke, & Berscheid, 1977; Todorov, Mandisodza, Goren, & Hall, 2005; Zebrowitz, 1996). Furthermore, photographs are an increasingly common feature of everyday life, making physical appearance accessible in many non-face-to-face contexts such as personal Web sites (Marcus, Machileik, & Schütz, 2006; Vazire & Gosling, 2004) and online social networks (e.g., Facebook; Buffardi & Campbell, 2008; Gosling, Gaddis, & Vazire, 2008). Despite the ubiquity of appearance and the prevalence of appearance-based judgments, psychological research reveals little about how personality is expressed in physical appearance and how appearance informs observers' judgments of personality.

In fact, after two decades of research based on zero-acquaintance judgments of personality, researchers still do not know the true magnitude of accuracy based on physical appearance alone. Virtually all zero-acquaintance

1661

studies are based on face-to-face interactions or videotaped behaviors (e.g., Albright, Kenny, & Malloy, 1988; Ambady & Rosenthal, 1992; Borkenau & Liebler, 1992; Carney, Colvin, & Hall, 2007; Funder & Sneed, 1993; Hall, Andrzejewski, Murphy, Mast, & Feinstein, 2008; Kenny, Horner, Kashy, & Chu, 1992; Levesque & Kenny, 1993; Paulhus & Bruce, 1992; Watson, 1989; Yeagley, Morling, & Nelson, 2007), which leaves room for an unknown quantity of supra-appearance information to be conveyed from targets to observers. Thus, the basic question remains unanswered: "How much can be learned about personality from physical appearance alone?"

Our study examines judgments at absolute zero-acquaintance—judgments based on physical appearance captured in full-body photographs. We ask which traits observers can accurately judge on the basis of physical appearance alone. Furthermore, because physical appearance has multiple components (e.g., facial expression, posture, clothing), we assess observer accuracy in two conditions: accuracy from a standardized, posed photograph that reduced the targets' nonverbal expressive behavior (*standardized condition*) and accuracy from a photograph in which targets were free to adopt a spontaneous pose and expression (*spontaneous condition*). Finally, using Brunswik's (1956) lens model as a framework, we examine whether specific static and dynamic appearance-based cues are associated with the targets' actual personalities (*cue validity*) and with the observers' judgments (*cue utilization*).

BACKGROUND

As we have noted, many zero-acquaintance findings are based on judgments made in information-rich settings such as face-to-face interactions or short video clips (Hall et al., 2008). Findings from these studies show that accuracy is often surprisingly high, especially for extraversion (Hall et al., 2008; Kenny, 1994). But we do not know how much of the accuracy achieved is due to sources of information over and above appearance. One study attempted to examine the different effects of physical appearance, verbal behavior, and nonverbal behavior on personality judgment (Borkenau & Liebler, 1992). Targets were videotaped entering a room, sitting behind a desk, and reading a weather forecast. The researchers presented observers with one of four stimuli: video with sound, video without sound, audio only, and a still extracted from the video. Observers who saw the full video with sound judged four of the Big Five traits accurately, whereas those who saw the video still judged only extraversion and conscientiousness accurately. These results suggest that physical appearance does offer some valid information, but accuracy increases when other verbal and nonverbal sources of information are available to observers.

Borkenau and Liebler's (1992) study was ground-breaking and informative, but its design limited the researchers' ability to estimate the magnitude of accuracy of judgments based on physical appearance alone. First, the targets in the still videos were sitting behind a desk, thereby obstructing observers' views of some aspects of physical appearance (e.g., clothing, posture). The results, then, may have underestimated both the accuracy levels and the range of traits that can be judged accurately based on physical appearance. Second, by restricting their examination to the Big Five personality traits, the researchers may have missed other important aspects of personality that can be judged on the basis of physical appearance. Third, accuracy was measured by correlating observers' ratings with a criterion measure based solely on self-reports. Multimethod criterion measures are more reliable and more valid than criteria from a single source (Funder, 1995; Kenny, 1994; Vazire, 2006), and self-reports are especially likely to be vulnerable to bias for some traits (e.g., evaluative, ambiguous traits; Dunning, Meyerowitz, & Holzberg, 1989; Vazire, 2008).

Other studies have examined the accuracy of personality judgments based on photographs alone, thus providing a test of accuracy at absolute zero acquaintance. Indeed, several studies have found substantial accuracy for some traits (e.g., Berry & Finch-Wero, 1993; Rind & Gaudet, 1993; Robins, Gosling, & Donahue, 1997; Vazire, Naumann, Rentfrow, & Gosling, 2008). However, these studies also had important design limitations. First, the bulk of this research has relied exclusively on facial photographs (see Zebrowitz & Collins, 1997, for a review) because the studies were designed to examine the role of facial expression in personality judgment, not the role of physical appearance in its entirety. As a result, many researchers used headshots or above-the-waist photographs as stimuli, not full-body photographs (e.g., Berry & Finch-Wero, 1993; Borkenau, Brecke, Möttig, & Paelecke, 2009; Little & Perrett, 2006; Penton-Voak, Pound, Little, & Perrett, 2006; Shevlin, Walker, Davies, Banyard, & Lewis, 2003; Zebrowitz, Hall, Murphy, & Rhodes, 2002). Knowing how facial expression contributes to judgments of personality is important, but meta-analyses of thin-slice judgments suggest that many cues to personality also reside below the head (e.g., clothing, posture; Ambady & Rosenthal, 1992). Another limitation of many of these studies is that they limit their examination of accuracy to just one or two traits, such as trustworthiness, narcissism, or intelligence (e.g., Todorov, Baron, & Oosterhof, 2008; Vazire et al., 2008; Zebrowitz et al., 2002). Finally, many of these studies used targets' self-ratings as the only accuracy criterion. Together, these design features may be underestimating

the magnitude of the accuracy correlations that can be achieved and the breadth of personality traits that can be accurately detected from physical appearance.

Finally, physical appearance is composed of multiple sources of information. It contains both static components related to physical grooming (e.g., style of dress and hairstyle) and dynamic aspects related to nonverbal expressive behavior (e.g., posture and facial expression; Riggio, Widaman, Tucker, & Salinas, 1991). Prior research has found that different components of physical appearance relate to different personality traits. For example, style of dress (e.g., formal attire) is a valid indicator of conscientiousness (Albright et al., 1988; Borkenau & Liebler, 1992) and facial expression has been connected to extraversion (Kenny et al., 1992). Furthermore, residue from expressive behavior may even pervade static aspects of appearance. For example, over a lifetime, emotionally expressive behavior can become etched on people's neutral faces (e.g., "laugh lines"; Malatesta, Fiore, & Messina, 1987). This pattern of evidence suggests that many elements of physical appearance may inform the process of accurate personality judgment, yet no research has compared the accuracy of judgments based on different components of physical appearance.

THE PRESENT STUDY

We examined three research questions: (a) What traits can be perceived accurately based on physical appearance in a standardized (posed) photograph? (b) Does accuracy improve when spontaneously expressed nonverbal components of physical appearance are visible to observers? and (c) Which static and dynamic appearance-based cues are associated with the targets' actual personalities (*cue validity*) and with the observers' judgments (*cue utilization*)? The answers to these questions will further elucidate the role of physical appearance and its multiple components in personality judgment.

We incorporated several critical design features that allowed us to provide a rigorous test of our research questions. First, to capture all aspects of physical appearance we used unobstructed full-body photographs of targets as stimuli. Second, because physical appearance is composed of multiple static and dynamic aspects, we attempted to systematically restrict the nonverbal expressive portion of physical appearance by taking two photographs of each target—one in which targets were told to stand in a standardized fashion with a neutral facial expression and posture (standardized condition) and one in which targets were given no instructions about how to pose (spontaneous condition). Third, we coded both static and dynamic appearance-based cues that may serve as the lens through which observers form

accurate judgments about targets' personalities. Fourth, we examined a broad range of personality traits—the Big Five traits, as well as likability, self-esteem, loneliness, religiosity, and political orientation. By including the Big Five traits, we were able to compare our findings to those of other zero-acquaintance research. By including five additional traits that are important in interpersonal perception but are not well captured by the Big Five, we were able to examine other potential domains of accuracy in personality judgment. Fifth, our accuracy criterion was based on a multi-method composite measure of self-ratings and ratings by well-acquainted informants. In the following we describe how we tested each research question.

Question 1: What traits can be perceived accurately based on physical appearance in a standardized photograph? To test this question, we examined the accuracy of judgments made in the standardized condition by correlating observers' ratings with our criterion measure of what the targets were actually like (i.e., the self-informant composite). Because of the dearth of research on non-Big-Five traits, we did not make specific predictions about how accuracy would vary across these traits. Among the Big Five traits, previous zero-acquaintance studies based on slices "thicker" than a photograph have demonstrated that observers can most consistently judge extraversion with some degree of accuracy (Borkenau & Liebler, 1992; Carney et al., 2007; Hall et al., 2008; Kenny, 1994). Therefore, we predicted that observers would be able to judge extraversion from a standardized photograph.

We also predicted that observers would be able to judge openness accurately. Accuracy is rarely achieved for openness when observers watch brief film clips of targets or engage in brief face-to-face interactions with targets (e.g., Albright et al., 1988; Funder & Colvin, 1988). Some suggest that this is because openness is unlikely to be related to expressive behavior (Ambady, Bernieri, & Richeson, 2000; Carney et al., 2007). Still, other zero-acquaintance research whose stimuli rely on an individuals' crafted environment (e.g., offices, bedrooms, Web sites; Gosling, Ko, Mannarelli, & Morris, 2002; Marcus et al., 2006; Vazire & Gosling, 2004) has found observer accuracy for openness. In these environments, individuals can make identity claims about their personality and beliefs (e.g., through posters or buttons making a statement), and these claims may act as clues to their openness. We suspect that clothing may also serve as an outlet for identity claims. Therefore, we predicted that observers would be able to judge openness accurately from a standardized photograph when the target's full body is available.

Question 2: Does accuracy improve when nonverbal expressive behavior is visible to observers? To test this question, we compared the accuracy levels of judgments

formed in the standardized condition with those made in the spontaneous condition. Prior "thin-slice" research that has parsed stimuli into various channels of information (e.g., face vs. full body; silent vs. full-sound film) typically finds that the greater the information, the greater the overall accuracy (e.g., Ambady & Rosenthal, 1992; Borkenau & Liebler, 1992). Furthermore, research shows that nonverbal expressive behavior is particularly informative of an individual's underlying dispositions and internal states (DePaulo, 1992; Ekman & Friesen, 1969, 1974). Therefore, because the spontaneous condition increased the amount of nonverbal expressive behavior available to observers, we predicted that judgments made in the spontaneous condition would have greater accuracy than judgments made in the standardized condition overall.

*Question 3: Which cues are associated with the targets' actual personalities (*cue validity*) and with the observers' judgments (*cue utilization*)?* Physical appearance is rife with symbols reflecting a person's identity (identity claims) and traces of past or anticipated behavior (behavior residue; Gosling et al., 2002). Although little research has examined the role of appearance-based cues in personality judgments, a handful of studies have shown that style of dress (e.g., formal attire) is a valid indicator of conscientiousness (Albright et al., 1988; Borkenau & Liebler, 1992) and a cheerful expression is predictive of extraversion (Borkenau et al., in press; Kenny et al., 1992). These findings suggest that observers use both static (e.g., clothing style) and dynamic (e.g., smiling) elements of a target's physical appearance to form impressions and that these cues are differentially related to various personality characteristics. Therefore, we coded cues that captured both static and dynamic components of appearance in the spontaneously expressed photograph.

Next, we used Brunswik's (1956) lens model as an organizing framework to further explicate the process by which observers use these appearance-based cues to form accurate personality impressions. According to this model, observers look to cues in the environment (the lens) to formulate impressions of a target's underlying disposition. Observers make accurate judgments if they use cues that are valid indicators (and ignore invalid cues) of the target's underlying personality. Using this model as a guide, we first examined which cues were valid indicators of the targets' personalities (*cue validity*). Next, we examined whether observers may have used these cues to formulate their personality judgments (*cue utilization*). We predicted that the targets' personalities would be related to appearance-based cues and that these cues would be related to observers' judgments.

METHOD

Target Participants

Target participants (those photographed and used as the targets of personality judgment) were 123 undergraduate students enrolled in introductory psychology classes at the University of Texas at Austin. Targets were 55% female, 56% White, 23% Asian American, 12% Latino or Latina, 3% African American, 3% Other, and averaged 18.7 years old (*SD* = 2.0). Participants completed the experiment in exchange for partial fulfillment of course requirements. This sample is the same as that used in Vazire et al. (2008).

Two photographs were taken of each target participant—first in the spontaneous condition (where participants were not instructed on how to stand or pose), followed by the standardized condition (where participants were instructed to look directly into the camera, keep a neutral facial expression, and stand with their feet shoulder-width apart and hands at their sides). The location of the camera and the participant were fixed so that the bottom of the frame was just below the participant's feet, ensuring that the entire body would be captured in the photograph. The targets did not know that they would be photographed before coming to the experiment.

All targets in the standardized condition (100%) stood with their hands at their side, 94% stood with their feet shoulder-width apart, 92% looked directly at the camera, and 12% smiled, suggesting strong compliance with the instructions. Those who did not comply with instructions in the standardized condition were dropped from those analyses. In contrast, the same targets had more varied poses in the spontaneous condition: Sixty percent stood with their hands at their side, 30% stood with their feet shoulder-width apart, 66% looked directly at the camera, and 66% smiled, suggesting that there were individual differences in nonverbal expressive behavior. The 10 targets who smiled in the standardized condition were excluded from all analyses, resulting in a final sample of 113 targets.

Accuracy Criterion

Our criterion measure was based on the aggregate of the targets' self-ratings and their informants' ratings of them, a criterion now widely regarded as the gold standard in personality research (Funder, 1995; Funder & Sneed, 1993; Vazire, 2006; Vazire & Gosling, 2004). We assessed the targets' personalities using the 44-item Big Five Inventory (BFI; John, Naumann, & Soto, 2008) and a variety of single-item measures, including likability, self-esteem, loneliness, religiosity, and political liberalism. All items were rated

on a 7-point Likert-type scale ranging from 1 (*disagree strongly*) to 7 (*agree strongly*).

The target individuals completed the questionnaire in the laboratory and nominated three informants who knew them well. The informants were contacted by e-mail and completed the questionnaire online. We assured the informants that their ratings would be kept confidential. We obtained ratings from at least one informant for all of the targets; 47 targets had ratings from two informants, and 70 targets had ratings from all three informants. Acquaintance between the informants and the targets averaged 9.6 years (*SD* = 7.3), and 261 of the 305 informants (86%) reported knowing the participant "quite well" or "very well."

To create the criterion measure, we first aggregated across the available informant reports (because targets could have had one to three informant reports) and then averaged the composite informant ratings with the targets' self-ratings; therefore, self-ratings were equally weighted (i.e., ½) with the informant ratings for all targets. The intraclass correlation (ICC[2, 1]; Shrout & Fleiss, 1979) reliability for the self-informant aggregates was strong (mean α = .79).

Observer Ratings

Observers independently rated the personalities of the targets on the BFI and the five single-item measures described previously. In the standardized condition, six observers completed their ratings after viewing the standardized photograph of each target. These observers were told that the targets had been directed to pose in the standardized position. In the spontaneous condition, six diffrent observers independently rated the personalities of the same targets after viewing the spontaneous photograph of each target.

The observers were undergraduate students working on the project as research apprentices. The order in which the observers viewed the photographs was varied. None of the observers knew the targets, and the observers were instructed not to communicate with each other about their ratings. The ICC reliabilities for the aggregate of the observers' ratings in the standardized and spontaneous conditions were acceptable (mean αs = .49 and .69, respectively).

Cues

Cue selection. We drew from the few studies that examined the link between personality and appearance-based cues to inform our cue selection. In particular, we were interested in coding the static (e.g., style of dress; body appearance) and dynamic (e.g., nonverbal expressive behavior) aspects of physical appearance. The static

cues we coded were: healthy (vs. sickly) appearance, stylish (vs. unstylish) appearance, distinctive (vs. ordinary) appearance, and neat (vs. messy) appearance. The dynamic cues we coded were: smiling, looking away from (vs. straight at) camera, arms folded, arms behind back, energetic (vs. tired) stance, and tense (vs. relaxed) stance.

Coders. Two-person teams were assigned to code one of the selected cues from the photograph in the spontaneous condition. These coders were undergraduate students who had not participated in any previous phase of the study. None of the coders knew any of the targets. Coders were told to code only the cue given and not to make broader inferences about the targets. The most concrete cues, such as smiling and arms folded, were coded for their *presence* (1) or *absence* (0). More global cues, such as energetic stance, were rated on a 7-point bipolar scale (e.g., 1 = *tired*, 7 = *energetic*). We aggregated the pair of coders' ratings to form a composite rating for each cue. Alpha reliabilities for the appearance-based cues were healthy appearance (.65), stylish appearance (.69), distinctive appearance (.70), neat appearance (.70), smiling (.91), looking away from camera (.74), arms folded (.87), arms behind back (.97), energetic stance (.81), and tense stance (.67).

RESULTS

Definition of Accuracy

In line with previous research (e.g., Ambady et al., 2000; Bernieri, 2001; Bernieri, Gillis, Davis, & Grahe, 1996; Carney et al., 2007; Hall, Bernieri, & Carney, 2005; Hall et al., 2008), we tested whether observers were accurate using two indices: aggregated-observer accuracy (also "group" or "pooled" accuracy) and single-observer accuracy. The aggregated-observer accuracy index reflects the level of accuracy achieved for a group of raters independent of the idiosyncrasies of any single rater (Block, 1961). Aggregated-observer accuracy also increases the reliability of judgments (compared to single-observer accuracy) through the use of multiple items (i.e., judges) and is one of the most commonly used indices of accuracy (Kenny, 1994). We computed aggregated-observer accuracy by correlating the aggregate of the six observers' ratings with the accuracy criterion (aggregated self and informant ratings) in both the standardized and spontaneous conditions. The single-observer accuracy index reflects how accurate a single observer would be when forming impressions of a set of targets; single-observer accuracy generalizes to how accurate a "typical" observer would be. We computed single-observer

accuracy by correlating each observer's rating with the accuracy criterion separately for each observer and for all 10 traits. To summarize these results, we averaged across the six observers' accuracy correlations for each of the 10 personality traits and in both the standardized and spontaneous conditions using Fisher's r to z formula, and then transformed the average back into a correlation.

How much accuracy should we expect? In a minimal-information context such as this one, it may be difficult to know how much accuracy is a lot. One standard to help readers evaluate the magnitude of the accuracy correlations is to compare them to accuracy results found in other contexts. For example, a review by Hall et al. (2008) found that the average accuracy (when converted to a Pearson's r) for Big Five trait judgments based on a dynamic stimulus (e.g., movie or face-to-face interaction) was .23, ranging from .40 for extraversion to .12 for agreeableness. Two studies (Borkenau et al., in press; Carney et al., 2007) looked at accuracy after exposure to very brief (50 ms and 5 s, respectively) videos of the targets. In the 50 ms condition (Borkenau et al., in press), accuracy averaged .13 across the Big Five and ranged from .31 for extraversion to .01 for neuroticism. In the 5 s condition (Carney et al., 2007), accuracy averaged .14 across the Big Five and ranged from .22 for extraversion to .04 for agreeableness.

Accuracy correlations can also be compared to typical effect sizes across all of social/personality research. A recent paper (Richard, Bond, & Stokes-Zoota, 2003) presented a meta-analysis and reported that the average effect size (when converted to a Pearson's r) is .21 for the field as a whole ($SD = .15$), and .21 ($SD = .14$) for personality research in particular. These values give us a context for interpreting the magnitude of the accuracy correlations obtained in the present study.

In addition, because the correlation metric is not familiar to all readers, we follow the lead of Hall et al. (2008) and converted this index using the binomial effect size display (BESD; Rosenthal & Rubin, 1982) formula to an index called a proportion index, or pi (Rosenthal & Rubin, 1989). This index is somewhat controversial (see Hall et al., 2008, for a discussion) but provides a rough estimate of the proportion of targets that observers would have judged accurately if the measure of accuracy were binary (correct/incorrect) rather than a correlation.

Question 1: What Traits Can Be Perceived Accurately Based on Physical Appearance in a Standardized Photograph?

We predicted that when shown the targets' posed physical appearance, observers would form accurate judgments of extraversion and openness. To test this prediction, we

computed correlations between the observers' ratings in the standardized condition and the accuracy criterion for each trait. Consistent with our predictions, aggregated observers' ratings showed some degree of accuracy for extraversion and marginally significant levels of accuracy for openness and emotional stability (see first data column of Table 1). However, observers were not accurate at judging agreeableness and conscientiousness. In addition, observers' judgments of self-esteem and religiosity were also accurate at above chance levels when observers' judgments were aggregated. When accuracy was examined for the average single observer, only judgments of extraversion were accurate above chance levels (second data column of Table 1).

Question 2: Does Accuracy Improve When Nonverbal Expressive Behavior Is Visible to Observers?

We next tested the hypothesis that expressive behavior (e.g., facial and postural expression) improves observer accuracy. We predicted that accuracy would generally be stronger in the spontaneous condition than in the standardized condition. To test this prediction, we first computed accuracy correlations for the spontaneous condition (third and fourth data columns of Table 1). Consistent with our prediction, more traits were judged accurately in the spontaneous condition (9 vs. 5 of 10) and the average accuracy correlation was higher (.25 vs. .14). Next, we tested whether the difference in accuracy between the two conditions was statistically significant for each trait using Hotelling's t test of significance of the difference between dependent correlations with the Williams modification (Kenny, Kashy, & Cook, 2006). These analyses show that the aggregated observers' ratings in the spontaneous condition were more accurate than the aggregated observers' ratings in the standardized condition for agreeableness and openness, and marginally more accurate for conscientiousness and loneliness.

Another way to examine whether observers in the spontaneous condition achieved greater accuracy than observers in the standardized condition is to examine whether the observers' ratings in the spontaneous condition provided incremental validity over the observers' ratings in the standardized condition in predicting the targets' actual personalities. To do this, we conducted hierarchical regression analyses predicting the targets' personalities (i.e., the accuracy criterion) from observations made in the standardized and spontaneous conditions for each of the 10 traits. In Step 1 we entered the observations made in the standardized condition, and in Step 2 we simultaneously entered observations from both conditions. The standardized regression weights from Step 2 and change in R are presented in Table 2. A significant

TABLE 1: Accuracy of Impressions Formed in the Standardized and Spontaneous Conditions

| | Standardized Condition | | | | Spontaneous Condition | | | |
| | Aggregated Observer | | Single Observer | | Aggregated Observer | | Single Observer | |
Personality Trait	r	(pi)	r	(pi)	r	(pi)	r	(pi)
Extraversion	.39**	(.70)	.29**	(.65)	.42**	(.71)	.34**	(.67)
Agreeableness	−.11	(.45)	−.05	(.48)	.20*	(.60)	.13	(.57)
Conscientiousness	−.03	(.49)	−.01	(.50)	.12	(.56)	.08	(.54)
Emotional stability	.17†	(.59)	.09	(.55)	.18†	(.59)	.10	(.55)
Openness	.17†	(.59)	.12	(.56)	.35**	(.68)	.20*	(.60)
Likability	.10	(.55)	.07	(.54)	.28**	(.64)	.19*	(.60)
Self-esteem	.26**	(.63)	.14	(.57)	.28**	(.64)	.20*	(.60)
Loneliness	.06	(.53)	.04	(.52)	.23*	(.62)	.16†	(.58)
Religiosity	.24**	(.62)	.11	(.56)	.27**	(.64)	.15†	(.58)
Political orientation	.16†	(.58)	.10	(.55)	.17†	(.59)	.10	(.55)
Mean	.14	(.57)	.09	(.55)	.25	(.63)	.17	(.59)

NOTE: Aggregated observer is the correlation between the aggregated observers' ratings and the accuracy criterion. Single observer is the mean of the six pairwise correlations between each observer's rating with the accuracy criterion. Numbers in parentheses are the proportion index (*pi*) which is the proportion of targets that observers would have judged accurately if the measure of accuracy were binary (correct/incorrect) rather than a correlation, calculated using the Binomial Effect Size Display (BESD) formula (Rosenthal & Rubin, 1982). Means were computed using Fisher's *r*-to-Z transformations.
†*p* < .10, two-tailed. **p* < .05, two-tailed. ***p* < .01, two-tailed.

TABLE 2: Does Accuracy Improve When Expressive Behavior Is Visible to Observers?

| | Standardized | Spontaneous | |
Personality Trait	β	β	ΔR
Extraversion	.24*	.29**	.07**
Agreeableness	−.18	.25**	.16**
Conscientiousness	−.17	.23	.15
Emotional stability	.13	.15	.05
Openness	−.02	.36**	.18**
Likability	.03	.27**	.18**
Self-esteem	.17	.19	.05
Loneliness	−.05	.25*	.18*
Religiosity	.15	.20*	.06*
Political orientation	.09	.12	.03
Mean	.04	.23	.11

NOTE: *N* = 113 targets. β is the standardized regression weight from Step 2 of the hierarchical regression model that controls for aggregated observations made in the other condition. Δ*R* is the increase in the multiple correlation obtained when+ observations made in the spontaneous condition are added to the regression equation in Step 2. Means were computed using Fisher's *r* to *Z* transformations.
*p < .05, two-tailed. ***p* < .01, two-tailed.

change in *R* suggests that information in the spontaneous condition contributes to increases in accuracy over and above the information in the standardized condition. As these results show, observers' ratings in the spontaneous condition provided incremental validity over the observers' ratings in the standardized condition for 6 of the 10 traits.

How accurate was the typical observer, individually, in the spontaneous condition? The accuracy correlations for the average single observer were significant for

extraversion, openness, likability, and self-esteem and marginally significant for loneliness and religiosity (fourth data column of Table 1). Again, more traits were judged accurately in the spontaneous condition (4 vs. 1 of 10) and the average accuracy correlation was higher (.17 vs. .09).

Question 3: Which Cues Are Associated With the Targets' Actual Personalities (Cue Validity) and With the Observers' Judgments (Cue Utilization)?

For each cue and for each trait, we tested (a) the extent to which the cues in the spontaneously posed photos correlated with the targets' actual personalities, as measured with the self-informant composite (cue validity) and (b) the extent to which the cues correlated with observers' personality ratings made in the spontaneous condition (cue utilization). To measure cue validity, we correlated the targets' criterion scores (the aggregate of self and informant personality ratings) with the independently coded appearance-based cues. To measure cue utilization, we correlated aggregated observers' ratings from the spontaneous condition with these cues.

Table 3 presents the cue validity and cue utilization correlations for the Big Five dimensions. Both static and dynamic appearance-based cues were valid indicators of extraversion; the most valid cues included having an energetic stance, looking stylish and healthy, and smiling. Observers' impressions of extraversion were most strongly correlated with targets' energetic stance and smiling. Agreeable targets were more likely to smile and stand in a more relaxed way; observers' judgments in the spontaneous condition were also correlated with these cues.

TABLE 3: A Brunswik (1956) Lens Model Analysis of Judgments Based on Appearance for the Big Five Dimensions

Cue Validity Correlations						Cue Utilization Correlations				
E	A	C	ES	O	Appearance-Based Cues	E	A	C	ES	O
					Static Cues					
.33**	.09	.14	.21*	−.20*	Healthy (vs. sickly) appearance	.30**	.22*	.25**	.35**	−.01
.35**	−.04	−.05	.13	−.09	Stylish (vs. unstylish) appearance	.14	.03	.01	.17	.00
−.08	.00	−.21*	−.05	.19*	Distinctive (vs. ordinary) appearance	−.06	.00	−.15	.06	.35**
.22*	.04	.18	.03	−.31**	Neat (vs. messy) appearance	.19*	.10	.53**	.06	−.04
					Dynamic Cues					
.28**	.21*	.15	.14	.02	Smiling	.71**	.62**	.40**	.54**	.47**
−.02	.03	.02	−.08	.23*	Looking away from camera	−.07	−.10	.01	−.11	.16
−.24**	.13	−.17	.08	.00	Arms folded	.01	.08	.02	.04	−.04
.15	.14	.13	.11	.01	Arms behind back	.17	.15	.02	.25**	.18
.39**	.13	.11	.15	.03	Energetic (vs. tired) stance	.76**	.56**	.33**	.53**	.42**
−.27**	−.26**	−.09	−.20*	−.01	Tense (vs. relaxed) stance	−.66**	−.50**	−.01	−.62**	−.29**

NOTE: E = extraversion; A = agreeableness; C = conscientiousness; ES = emotional stability; O = openness.
*p < .05, two-tailed. **p < .01, two-tailed.

TABLE 4: A Brunswik (1956) Lens Model Analysis of Judgments Based on Appearance for Non-Big-Five Personality Traits

Cue Validity Correlations						Cue Utilization Correlations				
LIK	SE	LON	REL	POL	Appearance-Based Cues	LIK	SE	LON	REL	POL
					Static Cues					
.21*	.33**	−.27**	.09	.03	Healthy (vs. sickly) appearance	.31**	.42**	−.38**	.17	−.21*
.14	.17	−.23**	−.04	−.01	Stylish (vs. unstylish) appearance	.17	.28**	−.16	.04	−.10
−.22*	−.11	.08	−.08	.09	Distinctive (vs. ordinary) appearance	.04	−.08	.12	−.20*	.41**
.17	.11	−.27**	.13	.13	Neat (vs. messy) appearance	.21*	.39**	−.27**	.44**	.40**
					Dynamic Cues					
.26**	.24**	−.15	.24**	.02	Smiling	.67**	.61**	−.62**	.34**	.02
.03	.07	−.06	−.18	.08	Looking away from camera	.06	.03	−.04	−.05	−.09
.01	−.13	−.04	−.03	−.02	Arms folded	.04	−.04	−.03	−.02	.04
.19*	.22*	−.15	−.01	−.06	Arms behind back	.16	.24**	−.18	.06	.09
.27**	.21*	−.33**	.21**	−.07	Energetic (vs. tired) stance	.70**	.65**	−.62**	.31**	−.02
−.23*	−.18	.29**	−.24**	.02	Tense (vs. relaxed) stance	−.55**	−.53**	.58**	−.09	−.01

NOTE: LIK = likability; SE = self-esteem; LON = loneliness; REL = religiosity; POL = political orientation (liberal).
*p < .05, two-tailed. **p < .01, two-tailed.

Conscientious targets were less likely to dress in distinctive ways. However, observers' judgments did not correlate with this cue; instead, observers' judgments correlated with targets looking neat, healthy, and energetic and smiling. Emotionally stable targets were healthier looking and stood in more relaxed ways; observers' judgments correlated with these cues but correlated with invalid cues such as whether the target was smiling and stood with an energetic stance. Targets who were more open to experience were less likely to have a healthy, neat appearance but were more likely to have a distinctive style of dress and look away from the camera. Observers' judgments correlated with targets' distinctive appearance but did not correlate with the other valid cues, instead correlating with targets' smiling and energetic stance.

Table 4 presents the cue validity and cue utilization correlations for the other non-Big-Five dimensions. The most likable targets smiled, looked healthier, and stood in an energetic manner, but they also looked less tense, distinctive, and stood with their arms behind their backs. Observers' judgments of likability correlated strongly with targets' facial and postural expression; their judgments correlated with targets' smiling, energetic stance, and relaxed stance. Targets who had higher self-esteem looked healthier, smiled more, had an energetic stance, and stood with their arms behind their backs. Observers' judgments correlated with both the dynamic facial and postural cues, as well as the static cues such as having a neat, stylish, and healthy looking appearance. The loneliest targets stood in less energetic

and more tense ways and appeared more sickly, messy, and unstylish. Observers' judgments correlated with most of these cues, particularly targets' negative postural expression, but also correlated with invalid cues such as smiling. Religious targets were more likely to smile and stand in energetic and relaxed ways; observers' judgments correlated with smiling and energetic stance, but also correlated with targets' neat appearance. None of the appearance-based cues we coded were valid indicators of the targets' liberal political orientation; in spite of this, observers' judgments correlated with appearance-based cues such as targets' distinctive, neat, and unhealthy appearance.

Target Gender Differences

Did accuracy levels differ if the target was male or female? We tested the difference between the aggregated-observer accuracy correlations made for male targets and female targets. There was only one significant difference. Observers' judgments of conscientiousness were more accurate for male targets, $r(57) = .31$, than for female targets, $r(56) = -.17$; $z = 2.55$, $p < .05$. We also examined the cue validity and cue utilization correlations by gender for conscientiousness. In males, actual conscientiousness was associated with a more neat appearance, $r(57) = .33$, $p < .01$; looking healthier, $r(57) = .29$, $p < .01$; and having a more distinctive appearance, $r(57) = .29$, $p < .01$, and observers' judgments of conscientiousness were correlated with a target's neat appearance, $r(57) = .30$, $p < .01$, and looking healthy, $r(57) = .53$, $p < .01$. Among the cues examined here, there were no valid indicators of conscientiousness in female targets.

DISCUSSION

Our results underscore the important role of physical appearance in personality judgment. As we predicted, physical appearance serves as a channel through which personality is manifested and observers discover many aspects of personality. Indeed, even when observers' judgments were based on just a full-body photograph in which expressive aspects of targets' appearance were virtually eliminated, these judgments were accurate for extraversion, emotional stability, openness, self-esteem, and religiosity. When observers saw targets in their spontaneous poses, they were able to judge 9 of 10 traits with some degree of accuracy, and for 4 of those traits, the observers were more accurate than in the standardized condition. In addition, for 6 of the 10 traits, judgments made in the spontaneous condition provided incremental validity in predicting the targets' actual

personalities above and beyond judgments made in the standardized condition. These findings suggest that observers' lay theories about personality and its manifestation in physical appearance have some kernel of truth to them, especially when targets can choose their own posture and facial expression.

These findings provide important insights into the personality judgment process. Previous zero-acquaintance research has shown that accuracy occurs even with minimal information (e.g., the mean accuracy correlation based on viewing a brief interaction or short segments of videotaped behavior is .23; Hall et al., 2008). Our findings suggest that much of this accuracy may happen before any face-to-face interaction occurs or any behavior is observed. That is, observers are able to form reasonably accurate impressions for a number of traits simply on the basis of physical appearance (e.g., aggregated-observer accuracy in the spontaneous condition averaged .25 across all 10 traits and averaged .26 across the Big Five traits only). Furthermore, the most informative components lay in an individual's nonverbal expressive appearance such as facial expression and posture.

The similarities and differences with patterns of accuracy found in previous research are also informative. Like research in many different contexts, including face-to-face interactions, we found that observers rated extraversion most accurately (Borkenau et al., in press; Carney et al., 2007; Funder & Colvin, 1988; Hall et al., 2008; Kenny, 1994). Extraversion is typically associated with positive emotional expression (Borkenau et al., in press; Kenny et al., 1992); however, we found that observers could judge extraversion accurately even when we restricted the targets' facial expression. Cue analyses suggested that both static and dynamic appearance-based cues reflected extraversion; as expected, extraverts were more likely to smile and stand energetically, but they were also more likely to wear stylish clothes, have a neat appearance, and look healthier. These findings support previous research demonstrating that extraversion is an observable, easy trait to judge (Funder & Colvin, 1988; John & Robins, 1993).

Contrary to previous research, however, we found that observers made accurate judgments for emotional stability. The use of photographs may have given observers more opportunity to focus on and scrutinize the individual differences in posture that might not otherwise be possible in a face-to-face interaction. For example, cue analyses revealed that emotionally stable targets stood in a more relaxed stance, whereas neurotic targets stood in a more rigid and tense manner, and observers seemed to use this information.

Also surprising was the fact that, contrary to some previous research (Albright et al., 1988; Borkenau & Liebler, 1992; Kenny, 1994) we did not find accuracy

for conscientiousness in either the standardized or spontaneous condition. Why was this the case? First, much prior research suggests that observers use style or presentation of dress (e.g., neatness or formalness of clothing) to judge conscientiousness (Albright et al., 1988; Borkenau & Liebler, 1992; Kenny et al., 1992). Because our sample of targets was culled from a college campus whose daily uniform is composed of one's favorite pair of jeans and T-shirt, we may have captured a restricted range of conscientiousness-related physical appearance cues compared to other studies that included a more diverse set of targets (e.g., Borkenau & Liebler, 1992). However, upon closer inspection, we did find that observers were accurate at judging conscientiousness in male targets only. As the cue analyses revealed, observers were attempting to use cues that, intuitively, seem related to conscientiousness. For example, observers rated targets who had a neat and healthy appearance as more conscientious. However, these cues were only valid for the male targets. It is possible that gender norms restricted the variability in the neatness of dress among the female targets; that is, females may be expected to always look put together and presentable whereas males may not be held to the same appearance-based standards.

Although the levels of accuracy in our study are surprisingly high considering the very limited information available to observers, we do not mean to suggest that observers can learn everything they need to know about a person based on a snapshot. Indeed, the accuracy of single observers was substantially lower than that of the aggregated observers' ratings. Thus, any single impression based on a snapshot is far from perfectly accurate. Furthermore, several studies have demonstrated that the accuracy of judgments increases over time and with greater acquaintanceship (Funder & Colvin, 1988; Kenny, Albright, Malloy, & Kashy, 1994; Paulhus & Bruce, 1992) and that different settings offer richer information with which to judge certain traits (e.g., bedrooms for conscientiousness, Web sites for openness).

One implication of our findings is that individuals may choose to alter their appearance in specific ways either to make identity claims or to shape others' impressions of them. For example, some aspects of our static physical appearance are clearly malleable (e.g., unconventional clothing or hairstyle) and even some expressive aspects of appearance may be within our control (e.g., smiling). However, it may be difficult to control other channels of information such as our posture (DePaulo, 1992; Ekman & Friesen, 1974); it is here where much valid and useful information about our personalities lay. As our cue analyses show, posture cues such as having an energetic or tense stance were valid indicators for 5 of the 10 traits we examined and observers seemed to recognize the value in this less controllable nonverbal information.

In summary, our study expands on and refines our understanding of personality judgment and the role of physical appearance in impression formation. By using full-body photographs and examining a broad range of traits, we identified domains of accuracy that have been overlooked, leading to the conclusion that physical appearance may play a more important role in personality judgment than previously thought. In addition, our comparison of judgments in the standardized and spontaneous conditions provides the first direct evidence that although some accuracy is possible even without nonverbal expressive aspects of appearance, more traits can be detected and judgments are generally more accurate when nonverbal expressive behavior is available. Overall, the level and breadth of accuracy achieved underline the pervasiveness of personality. As John Irving's narrator stated, "Things often are as they appear."

REFERENCES

Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology, 55,* 387-395.

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 201-272). San Diego, CA: Academic Press.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111,* 165-181.

Bernieri, F. J. (2001). Toward a taxonomy of interpersonal sensitivity. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 3-20). Mahwah, NJ: Lawrence Erlbaum.

Bernieri, F. J., Gillis, J. S., Davis, J. M., & Grahe, J. E. (1996). Dyad rapport and the accuracy of its judgment across situations: A lens model analysis. *Journal of Personality and Social Psychology, 71,* 110-129.

Berry, D. S., & Finch-Wero, J. L. (1993). Accuracy in face perception: A view from ecological psychology. *Journal of Personality, 61,* 497-521.

Block, J. (1961). *The Q-sort method in personality assessment and psychiatric research.* Springfield, IL: Charles C. Thomas.

Borkenau, P., Brecke, S., Möttig, C., & Paelecke, M. (2009). Extraversion is accurately perceived after a 50-ms exposure to a face. *Journal of Research in Personality, 43,* 702-705.

Borkenau, P., & Liebler, A. (1992). Trait inferences: Sources of validity at zero-acquaintance. *Journal of Personality and Social Psychology, 62,* 645-657.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments* (2nd ed.). Berkeley: University of California Press.

Buffardi, L. E., & Campbell, W. K. (2008). Narcissism and social networking Web sites. *Personality and Social Psychology Bulletin, 34,* 1303-1314.

Carney, D. R., Colvin, C. R., & Hall, J. A. (2007). A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality, 41,* 1054-1072.

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin, 11,* 203-243.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology, 57,* 1082-1090.

Efran, M. G. (1974). The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. *Journal of Research in Personality, 8*, 45-54.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and cues to deception. *Psychiatry, 32*, 88-106.

Ekman, P. & Friesen, W. V. (1974). Detecting deception from body or face. *Journal of Personality and Social Psychology, 29*, 288-298.

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

Funder, D. C., & Colvin, C. R. (1988). Friends and strangers: Acquaintanceship, agreement, and the accuracy of personality judgment. *Journal of Personality and Social Psychology, 55*, 149-158.

Funder, D. C., & Sneed, C. D. (1993). Behavioral manifestations of personality: An ecological approach to judgmental accuracy. *Journal of Personality and Social Psychology, 64*, 479-490.

Gosling, S. D., Gaddis, S., & Vazire, S. (2008). First impressions based on the environments we create and inhabit. In N. Ambady & J. J. Skowronski (Eds.), *First impressions* (pp. 334-356). New York: Guilford.

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

Hall, J. A., Andrzejewski, S. A., Murphy, N. A., Mast, M. S., & Feinstein, B. A. (2008). Accuracy of judging others' traits and states: Comparing mean levels across tests. *Journal of Research in Personality, 42*, 1476-1489.

Hall, J. A., Bernieri, F. J., & Carney, D. R. (2005). Nonverbal behavior and interpersonal sensitivity. In J. A. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 237-281). New York: Oxford University Press.

John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big-Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114-158). New York: Guilford.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective on the self. *Journal of Personality, 61*, 521-551.

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

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

Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. (1992). Consensus at zero-acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology, 62*, 88-97.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis.* New York: Guilford.

Levesque, M. J., & Kenny, D. A. (1993). Accuracy of behavioral predictions at zero acquaintance: A social relations analysis. *Journal of Personality and Social Psychology, 65,* 1178-1187.

Little, A. C., & Perrett, D. I. (2006). Using composite images to assess accuracy in personality attribution to faces. *British Journal of Psychology, 98*, 111-126.

Malatesta, C. Z., Fiore, M. J., & Messina, J. J. (1987). Affect, personality, and facial expression characteristics of older people. *Psychology and Aging, 2,* 64-69.

Marcus, B., Machilek, F., & Schütz, A. (2006). Personality in cyberspace: Personal websites as media for personality expressions and impressions. *Journal of Personality and Social Psychology, 90,* 1014-1031.

Paulhus, D. L., & Bruce, M. N. (1992). The effect of acquaintanceship on the validity of personality impressions: A longitudinal study. *Journal of Personality and Social Psychology, 63,* 816-824.

Penton-Voak, I. S., Pound, N., Little, A. C., & Perrett, D.I. (2006). Personality judgments from natural and composite facial images: More evidence for a "kernel of truth" in social perception. *Social Cognition, 24*, 490-524.

Richard, F. D., Bond, C. F., Jr., & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology, 7,* 331-363.

Riggio, R. E., Widaman, K. F., Tucker, J. S., & Salinas, C. (1991). Beauty if more than skin deep: Components of attractiveness. *Basic and Applied Social Psychology, 12*, 423-439.

Rind, B., & Gaudet, S. (1993). Judging personality traits of adolescents from photographs. *Journal of Social Psychology, 133*, 815-819.

Robins, R. W., Gosling, S. D., & Donahue, E. M. (1997). Are personality judgments based on physical appearance consensual and accurate? In J. Bermúdez, B. de Raad, J. de Vries, A. M. Pérez-García, A. Sánchez-Elvira, & G. L. van Heck (Eds.), *Personality psychology in Europe* (Vol. 6, pp. 70-75). Tilburg, Netherlands: Tilburg University Press.

Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74*, 166-169.

Rosenthal, R., & Rubin, D. B. (1989). Effect size estimation for one-sample multiple-choice-type data. *Psychological Bulletin, 106*, 332-337.

Shevlin, M., Walker, S., Davies, M. N. O., Banyard, P., & Lewis, C. A., (2003). Can you judge a book by its cover? Evidence of self-stranger agreement on personality at zero acquaintance. *Personality and Individual Differences*, *35*, 1373-1383.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420-427.

Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology, 35,* 656-666.

Todorov, A., Baron, S., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: A model based approach. *Social, Cognitive, & Affective Neuroscience, 3*, 119-127.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308*, 1623-1626.

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality, 40,* 472-481.

Vazire, S. (2008). *Who knows what about a person? The self-other knowledge asymmetry (SOKA) model.* Manuscript under reciew.

Vazire, S., & Gosling, S. D. (2004). e-Perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology, 87,* 123-132.

Vazire, S., Naumann, L. P., Rentfrow, P. J., & Gosling, S. D. (2008). Portrait of a narcissist: Manifestations of narcissism in physical appearance. *Journal of Research in Personality*, *42*, 1439-1447.

Watson, D. (1989). Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report. *Journal of Personality and Social Psychology, 57,* 120-128.

Yeagley, E., Morling, B., & Nelson, M. (2007). Non-verbal zero-acquaintance accuracy of self-esteem, social dominance orientation, and satisfaction with life. *Journal of Research in Personality, 41,* 1099-1106.

Zebrowitz, L. A. (1996). Physical appearance as a basis of stereotyping. In C. N. Macrae, C. Stangor, & M. Hewstone (Eds.), *Stereotypes and stereotyping* (pp. 79-121). New York: Guilford.

Zebrowitz, L. A., & Collins, M. A. (1997). Accurate social perception at zero acquaintance: The affordances of a Gibsonian approach. *Personality and Social Psychology Review, 1*, 203-222.

Zebrowitz, L. A., Hall, J. A., Murphy, N. A., & Rhodes, G. (2002). Looking smart and looking good: Facial cues to intelligence and their origins. *Personality and Social Psychology Bulletin, 28*, 238-249.