Title Page (including author names and affiliations)

# Open Laptops, Open Minds:

# Consensus and Accuracy in Big Five Personality Perception from Laptop Stickers

Jessica T. Campbell,[1] Imani N. Turner,[2] and Gregory D. Webster[2]

[1]The Kinsey Institute, Indiana University, Bloomington, Indiana 47405 USA

[2]Department of Psychology, University of Florida, Gainesville, Florida 32611-2250 USA

December 17, 2021

**Manuscript under review; cite at own risk**

All-inclusive word count: ≈ 7,000

Author Note

Corresponding author: Gregory D. Webster, Department of Psychology, University of Florida, P.O. Box 112250, Gainesville, FL 32611-2250 USA. Email: gdwebs@ufl.edu

Jessica T. Campbell and Imani N. Turner contributed to developing the study idea, data collection and cleaning, and manuscript writing and editing; Gregory D. Webster contributed to data analysis, and manuscript writing and editing. This study was not preregistered. Data and code are available via OSF: https://osf.io/4ygca

The authors thank the following research assistants for their contributions to this work: El Rim Ahn, Jason Brown, Valentina Chapman, Madison Childers, Brittany Gendler, Madison Kohler, Tram Lai, Chris Lee, Luís Leme, Alexa Miller, Megan Mufson, Tanvi Padalkar, Roshni Patel, Rachel Walton, and Simon Zheng.

## Highlights

- Examined people's abilities to judge others' traits from their laptop stickers
- Compared traditional methods to cross-classified structural equation models (CC-SEMs)
- All Big Five traits except negative emotionality showed adequate consensus
- CC-SEM-based accuracy was significant only for extraversion and open-mindedness
- Number of stickers served as a cue for the accuracy effect of aesthetic sensitivity

## Abstract

Using traditional methods and cross-classified structural equation models (CC-SEMs), we assessed consensus and accuracy across 1,139 observations made by 8 perceivers of 147 laptops whose owners completed self-reports of the same Big Five traits and facets. Average Big Five consensus correlations were slightly stronger using traditional methods (.32) than using CC-SEMs (.25). Average accuracy correlations were slightly stronger using CC-SEMs (.21) than traditional methods (.18). Using CC-SEMs, all traits except negative emotionality showed adequate consensus, whereas accuracy was significant only for open-mindedness and extraversion; extraversion emerged only after controlling for laptop type and number-of-stickers. Number-of-stickers partially mediated—or served as a cue for—the accuracy effect linking targets' self-reports of aesthetic sensitivity to perceivers' ratings of the same open-mindedness facet.

**Keywords:** Big Five, Personality, Perception, Judgment, Consensus, Accuracy, Cross-Classification Analysis, Structural Equation Model, Narcissism

Open Laptops, Open Minds:

Consensus and Accuracy in Big Five Personality Perception from Laptop Stickers

# 1 Introduction

College students in the U.S. routinely adorn their laptop computers with decorative stickers. When used in public settings—classrooms, libraries, coffee shops—their opened laptops become veritable billboards projecting their interests and identities toward peers. Because these stickers may serve to communicate one's interests and shape impression formation, they are an ideal candidate cue for studying interpersonal perception. In this work, we examine the extent to which eight college students show consensus and accuracy when asked to rate 147 other students' personality traits and facets based solely on 147 photos of their laptop stickers. In this context, *consensus* reflects the amount of agreement among the set of eight perceivers (or raters) that attempts to infer personality from a set of target stimuli (people's laptops), whereas *accuracy* reflects the association between perceivers' perceptions of targets' personality traits and a criterion measure of personality (targets' self-reports).

## *1.1 Personality Perception: Perspectives and New Methods*

Personality perception or judgement has a long history in psychology (for reviews see Biesanz & Wallace, 2020; Funder, 1995; Hall et al., 2016). Most studies have focused on people's ability to infer strangers' personality traits in face-to-face settings at zero-acquaintance (Nestler & Back, 2013). Other studies have focused on behavioral residue or identity claims, such as those that can be inferred by viewing people's bedrooms and offices (Gosling et al., 2002), Facebook pages (Buffardi & Campbell, 2008), or even online-gaming avatars and usernames (Harari et al., 2015). Still other studies have focused on inferring preferences, such as music (Nave et al., 2018), clothing (Naumann et al., 2009), and shoes (Gillath et al., 2012). Because the stickers people choose to adorn their laptops carry elements of identity claims (e.g., likes biking, kayaking) and preferences (e.g., supports socialism, animal rights), the present study draws on both traditions.

The amount of consensus and accuracy present varies depending on the trait assessed and the person or object observed. For bedrooms/offices (Gosling et al., 2002), consensus correlations were highest for openness (.58/.51) and lowest for neuroticism (.08/.14), whereas accuracy correlations were highest for openness (.65/.46), and lowest for agreeableness (.20/-.04). For online social networks, a meta-analysis (Tskhay & Rule, 2014) showed that consensus and accuracy correlations were highest for extraversion (.32, .37) and lowest for neuroticism (.15, .08).

Methods used to assess consensus and accuracy differ based on theoretical approach. Most studies of preferences, residual behavioral, and identity claims use a lens-model approach (Brunswik, 1956), where various cues are used (e.g., organized desk, many books) by perceivers to infer targets' personalities. In such studies, consensus is a simple measure of inter-rater agreement, such as the mean inter-rater correlation (MIC), and accuracy is assessed by taking an average rating across all perceivers and correlating those averages with a criterion measure, often self- and/or peer reports (e.g., Gosling et al., 2002). Target cues are often assessed to help explain how accuracy is inferred. For example, a room's tidiness might be a cue utilized to infer conscientiousness. Lens models are tested as mediation models, where perceptions are regressed onto the criterion (paradoxically the focal predictor) and the cue (mediator).

Another approach to measuring interpersonal consensus and accuracy relies on partitioning variance in observers' ratings into target, perceiver, and residual variance

(Kenny, 2020). These variances can be used to (a) calculate measures of consensus (e.g., intraclass correlation coefficients [ICCs]) and (b) provide a purer measure—one that accounts for non-independence across perceivers instead of averaging across them—for examining associations with criterion measures to assess accuracy.

Building on both approaches, researchers have pioneered cross-classification models of accuracy, whereby both targets and perceivers are treated as random variables and the non-independences among both targets and perceivers is accounted for (Claus et al., 2020; Nestler & Back, 2017). Such models, which often use Bayesian estimation procedures, allow for not only variance partitioning, but also cue utilization in a mediation model, thereby combining powerful explanatory frameworks from both approaches. Further still, such cross-classified models can incorporate latent-variable modeling (structural equation modeling [SEM]), meaning that measurement error is modeled when one uses multiple measures of—or items for—the same trait (CC-SEM; Nestler & Back, 2017). To these ends, the present work first presents results using "traditional methods," which ignore perceiver-level non-independence and measurement error, followed by CC-SEM results, which account for both the cross-nested data structure and measurement error.

Regarding measures, the present study also sought to break new ground by assessing personality facets in addition to traits. Whereas traits reflect broad personality constructs (e.g., extraversion), facets focus on more specific aspects of personality (e.g., sociability; Soto & John, 2017). Many prior studies of personality perception have used either 10-item (e.g., TIPI; Gosling et al., 2003) or 44-item (BFI; John & Srivastava, 1999) Big Five measure, neither of which assesses facets.

### *1.2 Predictions*

We developed three hypotheses for both the Big Five and narcissism:

1. Big Five consensus coefficients will be at least small-to-moderate ($r$s ≈ .10 to .30), regardless of method. This prediction is based on a meta-analysis ($k$ = 18–19 studies) of Big Five personality consensus judgments from online social network profiles (Tskhay & Rule, 2014; $r$s .11 to .32, $M_r$ = .22).
2. Based on prior research on Big Five personality accuracy correlations for people's offices and bedrooms (Gosling et al., 2002), we expected accuracy correlations to be moderate-to-large for openness/open-mindedness ($r$s = .46 and .65, $M_r$ = .56) and small-to-moderate for other traits ($r$s = -.04 to .36, $M_r$ = .22).
3. We chose to assess narcissism because we believed that it would be reliably detectable from laptop stickers, and because of its inclusion in related research on people's perceptions of narcissism from others' social media profiles and activity (McCain & Campbell, 2018). Specifically, we expected small-to-moderate accuracy correlations based on prior research of narcissism judgments from people's Facebook pages (Buffardi & Campbell, 2008; $r$ = .25); however, because consensus correlations were not assessed, we had no specific predictions other than expecting narcissism consensus coefficients to be similar to Big Five ones.

Although we made no a priori attempt to assess cues (cf. Gosling et al., 2002), we did measure two variables that may have served as cues—laptop type and number of stickers—in aiding personality inference via laptops. Thus, as an exploratory exercise, we controlled for these cues as covariates in some analyses, and tested for lens-model mediation (or cue utilization).

## 2 Method

### *2.1 Participants and Procedure*

### *2.1.1 Targets: Laptops and Their Owners*

We recruited 172 students from a large public university in the southeastern U.S. enrolled in general psychology courses who had laptops with stickers and agreed to have them photographed in exchange for course credit. After giving their informed consent, a research assistant photographed the top panel of participants' laptops. Participants then completed demographic and personality items online using a laboratory computer. We collected data from June 2018 through March 2019. Target sample size was determined by examining similar studies of personality judgments of offices ($N$ = 94) and bedrooms ($N$ = 83; Gosling et al., 2002), physical appearance ($N$ = 113; Naumann et al., 2009), and a recent 28-study meta-analysis of online social network profiles ($M_N$ = 132), which together yielded a weighted mean sample size of 129 targets. To be conservative, we aimed to exceed this average by one-third or 172 targets.

We had two research assistants independently rate photos of laptops regarding their type—Mac, PC, or unknown (logos obscured)—and the number of stickers displayed. We assessed inter-rater reliability via simple correlation, which was .82 for laptop type and .87 for number of stickers. Rather than resolve discrepancies between raters based on ambiguities (e.g., Should a single two-piece sticker count as one or two stickers?), we averaged across raters resulting in a mean number of stickers and an index of "Mac-ness" ranging from 0 (definitely a PC) to 1 (definitely a Mac) in increments of 0.25 ($M$ = 0.70, $SD$ = 0.40). We excluded laptops with fewer than three stickers early on in determining our stimulus set and thus do not have complete ratings from laptops with only one or two stickers, yielding a target sample of 147. We made this admittedly arbitrary choice because we felt that three stickers were the minimum number that a person might need to view to form an accurate impression of another person's personality, in much the same way that three items are often used as the minimum number to accurately assess a psychological construct. Among these 147 laptops, the number of stickers ranged from 3 to 39 and was extremely positively skewed (skewness = 1.85, kurtosis = 4.30; $M$ = 10.46, $SD$ = 6.45); a natural-log transformation normalized the frequency distribution for analysis (skewness = 0.04, kurtosis = -0.02; $M$ = 2.19, $SD$ = 0.56).

The analyzed sample of 147 participants was 82% women (119 cisgender women, 1 transgender woman, 24 cisgender men, 1 transgender man, 2 people who preferred not to respond), 65% White (95 Whites/Caucasians, 26 non-White Hispanics/Latinx, 16 Asians/Pacific Islanders, 6 Blacks/African-Americans, 4 people specifying another race/ethnicity or declining to respond), and young, ranging in age from 18 to 24 years ($M$ = 19.05, $SD$ = 1.14).

### *2.1.2 Perceivers*

We recruited eight undergraduate research assistants (7 women, 1 man) from two psychology labs. These eight perceivers were asked to rate the 147 digital photos of the targets' laptops on the Big Five traits, their facets, and narcissism.

### *2.1.3 Power*

Because two perceivers did not rate all targets, the total sample size was 1,139 observations (97% of a possible 1,176) cross-classified within 147 targets and eight perceivers. Determining precise statistical power for such designs is complex; however, because we sought to generalize mainly at the target level (participants and their laptops), we based our power analysis on our sample of 147, which yielded adequate power (≥ .80) to detected effect sizes of $|r| \geq .23$ ($R^2 \geq .05$), assuming $\alpha = .05$. This is just above the threshold for the average Big Five consensus ($M_r$ = .22) and accuracy ($M_r$

= .22) correlations observed in a nearly 20-study meta-analysis of personality perception based on online social network profiles (Tskhay & Rule, 2014).

### *2.2 Measures*

Both targets and perceivers used the same measures, which used response scales from 1 (*Disagree strongly*) to 5 (*Agree strongly*). All false-keyed items were reverse-scored prior to analyses, with higher scores reflecting more of each trait or facet.

#### *2.2.1 Big Five Personality*

We assessed personality traits and facets using the 60-item Big Five Inventory 2 (BFI-2; Soto & John, 2017), which contains 3 four-item facets for each Big Five trait (15 facets; see Table 1). Items included "…is dominant, acts as a lead" for extraversion (assertiveness) and "…is fascinated by art, music, or literature" for open-mindedness (aesthetic sensitivity). Cronbach's alphas were good for rated (≥ .82) and self-reported ≥ .80) 12-item traits, and generally acceptable for rated (range: .54 to .86; $M$ = .73, $SD$ = .09) and self-reported (range: .63 to .88; $M$ = .75, $SD$ = .07) four-item facets (Table 1).

#### *2.2.2 Narcissism*

We measured narcissism using all four items from the Dark Triad Dirty Dozen (Jonason & Webster, 2010); neither psychopathy nor Machiavellianism was assessed. Items included "I tend to want others to admire me" and "I tend to want others to pay attention to me." Cronbach's alphas were acceptable both rated (.82) and self-reported (.71) narcissism.

### *2.3 Data Analysis*

We analyzed data using R and Mplus (Muthén & Muthén, 2017). For traditional analyses (e.g., Gosling et al., 2002), we examined mean inter-rater correlations (MICs) for consensus and two types of correlations for accuracy. First, we assessed aggregated-

**Table 1.** *Descriptive Statistics for Targets' Self-Reports and Perceivers' Average Ratings*

| | Targets' Self-Reports | | | Perceivers' Average Ratings of Targets | | |
|---|---|---|---|---|---|---|
| Personality trait/facet | Mean | *SD* | α | Mean | *SD* | α |
| Extraversion | 3.57 | 0.71 | .87 | 3.44 | 0.50 | .88 |
| Sociability | 3.39 | 1.03 | .88 | 3.47 | 0.59 | .86 |
| Assertiveness | 3.42 | 0.88 | .80 | 3.06 | 0.56 | .74 |
| Energy | 3.91 | 0.75 | .71 | 3.80 | 0.51 | .78 |
| Agreeableness | 3.91 | 0.58 | .80 | 3.78 | 0.35 | .89 |
| Compassion | 4.13 | 0.75 | .63 | 3.87 | 0.38 | .79 |
| Respectfulness | 4.15 | 0.66 | .69 | 3.84 | 0.38 | .77 |
| Trust | 3.45 | 0.79 | .67 | 3.63 | 0.37 | .65 |
| Conscientiousness | 3.77 | 0.65 | .86 | 3.43 | 0.39 | .84 |
| Organization | 3.87 | 0.90 | .84 | 3.36 | 0.57 | .79 |
| Productiveness | 3.65 | 0.77 | .73 | 3.54 | 0.39 | .75 |
| Responsibility | 3.77 | 0.70 | .70 | 3.40 | 0.35 | .57 |
| Negative emotionality | 3.03 | 0.81 | .90 | 2.55 | 0.22 | .82 |
| Anxiety | 3.56 | 0.88 | .76 | 2.68 | 0.26 | .54 |
| Depression | 2.66 | 0.95 | .82 | 2.32 | 0.29 | .74 |
| Emotional volatility | 2.88 | 0.99 | .85 | 2.66 | 0.27 | .67 |
| Open-mindedness | 3.89 | 0.66 | .85 | 3.44 | 0.51 | .89 |
| Aesthetic sensitivity | 3.70 | 0.94 | .75 | 3.30 | 0.61 | .80 |
| Intellectual curiosity | 4.06 | 0.73 | .74 | 3.45 | 0.49 | .71 |
| Creative imagination | 3.92 | 0.73 | .70 | 3.57 | 0.50 | .78 |
| Narcissism | 3.20 | 0.82 | .71 | 2.77 | 0.55 | .82 |

*Note*. *N*s = 1,139 observations of 147 participants and their laptops (targets) from 8 raters (perceivers).

observer correlations between self-reports and personality ratings averaged across perceivers for accuracy. Second, we assessed single-observer correlations, in which each perceiver's individual accuracy correlation with targets' self-reports are taken and then averaged across perceivers (see Naumann et al., 2009). For these traditional methods, we use correlations with 95% confidence intervals.

For advanced analyses, we used CC-SEMs (Nestler & Back, 2017) to estimate variance components, and hence consensus using ICCs (Claus et al., 2020; Kenny, 2020). Because these consensus ICCs are based on precise proportions of variance explained, they can be interpreted as correlations but have no corresponding confidence or credibility intervals. We assessed accuracy by regressing latent trait ratings from perceivers onto their respective latent trait ratings from targets via self-reports. We present both standardized (β) and unstandardized (*b*) regression coefficients with 95% Bayesian credibility intervals for CC-SEMs. We also examined Mac-ness and log number of laptop stickers as covariates or possible cues. Although we did not preregister this study, data and code (R and Mplus) for all analyses are available here: https://osf.io/4ygca/?view_only=47f4639812d3431282a444fd4ace5fd9

# 3 Results

## *3.1 Traditional Methods*

### *3.1.1 Consensus*

Big Five consensus correlations (MICs) were moderate for extraversion (.44), open-mindedness (.40), and conscientiousness (.39); moderate for agreeableness (.28); and small for negative emotionality (.09); their mean was .32 (*SD* = .14). Facet consensus correlations ranged from .06 to .38 with an average MIC of .26 (*SD* = .11; see Table 2, left column). The MIC for narcissism was .35.

**Table 2.** *Traditional Methods: Consensus and Accuracy Correlations*

| Personality trait/facet | Consensus | Accuracy | |
|---|---|---|---|
| | | Aggregated Observer | Single Observer |
| Extraversion | .44 [.29, .56] | .15 [-.01, .31] | .10 [-.07, .25] |
| Sociability | .37 [.22, .50] | .11 [-.06, .26] | .07 [-.10, .23] |
| Assertiveness | .34 [.18, .48] | .18 [.01, .33] | .11 [-.06, .27] |
| Energy level | .38 [.23, .51] | .10 [-.07, .26] | .06 [-.10, .22] |
| Agreeableness | .28 [.12, .42] | .18 [.02, .33] | .13 [-.04, .28] |
| Compassion | .23 [.07, .38] | .09 [-.07, .25] | .06 [-.11, .22] |
| Respectfulness | .26 [.10, 40] | .16 [-.00, .31] | .10 [-.06, .26] |
| Trust | .20 [.04, .35] | .16 [-.00, .31] | .11 [-.05, .27] |
| Conscientiousness | .39 [.24, .52] | .17 [.01, .33] | .12 [-.04, .28] |
| Organization | .38 [.23, .51] | .15 [-.02, .30] | .11 [-.06, .26] |
| Productiveness | .25 [.09, .40] | .11 [-.05, .27] | .08 [-.08, .24] |
| Responsibility | .26 [.10, .40] | .08 [-.08, .24] | .06 [-.10, .22] |
| Negative emotionality | .09 [-.07, .24] | .10 [-.06, .26] | .06 [-.11, .22] |
| Anxiety | .06 [-.10, .22] | -.03 [-.19, .13] | -.01 [-.18, .15] |
| Depression | .13 [-.03, .29] | .08 [-.09, .23] | .04 [-.12, .20] |
| Emotional volatility | .08 [-.08, 24] | .20 [.04, .35] | .09 [-.06, .25] |
| Open-mindedness | .40 [.25, .53] | .30 [.14, .44] | .22 [.06, .37] |
| Aesthetic sensitivity | .36 [.21, .49] | .30 [.15, .44] | .20 [.04, .35] |
| Intellectual curiosity | .31 [.16, .45] | .28 [.13, .42] | .19 [.03, .34] |
| Creative imagination | .30 [.15, .44] | .11 [-.05, .27] | .09 [-.08, .25] |
| Narcissism | .35 [.20, .48] | .07 [-.09, .23] | .04 [-.12, .20] |
| Big Five trait mean | .32 [.16, 46] | .18 [.02, .33] | .12 [-.03, .28] |
| Big Five facet mean | .26 [.10, .40] | .14 [-.02, .30] | .09 [-.07, .25] |

*Note*. $N$ = 1,139 observations (147 persons and their laptops by 8 independent observers). 95% confidence intervals based on $N$ = 147.

*3.1.2 Accuracy*

Following prior work (Naumann et al., 2009), we assessed accuracy using two methods. The first used aggregated-observer ratings (see also Gosling et al, 2002), which simply takes the mean across all eight raters or perceivers for a given trait and correlates that average with targets' self-reports for the same trait. Big Five accuracy correlations were moderate for open-mindedness (.30 [.14, .44]); small-to-moderate for agreeableness (.18 [.02, .33]), conscientiousness (.17 [.01, .33]), and extraversion (.15 [-.01, .31]); and small for negative emotionality (.10 [-.06, .26]); their mean was .18 [.02, .33] (Table 2, middle column). Facet accuracy correlations ranged from -.03 to .30 with an average of .14 [-.02, .30] (Table 2, middle column). The accuracy correlation for narcissism was .07 [-.09, .23].

The second method used single-observer ratings, which takes a more idiographic approach by examining the correlations between each individual perceiver's rating of a given trait with targets' self-reports for the same trait. The resulting eight correlations (one for each perceiver) are then Fisher's *r*-to-*z*-transformed, averaged, and then back-transformed into an average correlation. Single-observer accuracy correlations are uniformly weaker than aggregate-observer accuracy correlations (see Naumann et al., 2009). Big Five accuracy correlations were small-to-moderate for open-mindedness (.22 [.06, .37]); small for agreeableness (.13 [-.04, .28]), conscientiousness (.12 [-.04, .28]), and extraversion (.10 [-.07, .25]); and trivial for negative emotionality (.06 [-.11, .22]); their mean was .12 [-.03, 28] (Table 2, right column). Facet accuracy correlations ranged from -.03 to .30 with an average of .09 [-.07, .25] (Table 2, right column). The accuracy correlation for narcissism was .04 [-.12, .20]. In contrast to aggregated-observer ratings, single-observer accuracy correlations were significant only for open-mindedness and two of its facets—aesthetic sensitivity and intellectual curiosity.

### ***3.2. Cross-Classified Structural Equation Models***

We specified a series of CC-SEMs (Nestler & Back, 2017) that allowed us to account for non-independent observations at both the rater and participant level. Specifically, we partitioned variance attributable to participants (targets), raters (perceivers), and observations (residual). To account for measurement error, we created latent variables at all three levels as well as participants' self-reports. For narcissism and each BFI-2 facet, four items loaded onto a latent factor. For the Big Five traits, we specified hierarchical or second-order latent measurement models, where individual items loaded onto their respective latent facet, and the three latent facets loaded onto their respective Big Five trait. For parsimony and to aid model convergence, identical items at different levels were constrained to have the same loadings. Thus, separating measurement error from true-score variance and partitioning variance by accounting for non-independence allowed for a purer test of accuracy, which we tested as the association between latent target ratings and latent self-reports of the same trait or facet.

#### *3.2.1 Consensus*

We assessed consensus by partitioning variance among target, perceiver, and residual sources and then calculating intraclass correlations (ICCs; i.e., target variance ÷ total variance; see Kenny, 2020; see Table 3, left column, boldfaced). Trait consensus ICCs ranged from moderate for extraversion (.39) to trivial for negative emotionality (.06) with a small-to-moderate mean of .25 ($SD$ = .14). Facet consensus ICCs similarly ranged from moderate for organization (.38) to trivial for anxiety (.05) with a small-to-

**Table 3.** *Variance Estimates [with 95% Bayesian credibility intervals] and Consensus Correlations (Boldface, Target ICCs) via Variance Decomposition Using Cross-Classified Structural Equation Models (CC-SEMs)*

| | Variance by source (proportions or ICCs) | | | | | |
|---|---|---|---|---|---|---|
| Personality trait/facet | Target [95% CI] | ICC | Perceiver [95% CI] | ICC | Residual [95% CI] | ICC |
| Extraversion | 0.252 [0.185, 0.343] | **.39** | 0.062 [0.007, 0.391] | .09 | 0.339 [0.289, 0.395] | .52 |
| Sociability | 0.236 [0.173, 0.325] | **.36** | 0.052 [0.010. 0.265] | .08 | 0.369 [0.321, 0.424] | .56 |
| Assertiveness | 0.255 [0.180, 0.360] | **.36** | 0.067 [0.002, 0.442] | .09 | 0.385 [0.315, 0.451] | .54 |
| Energy level | 0.158 [0.116, 0.220] | **.37** | 0.032 [0.003, 0.188] | .08 | 0.232 [0.196, 0.273] | .55 |
| Agreeableness | 0.095 [0.066, 0.136] | **.17** | 0.250 [0.075, 1.168] | .44 | 0.225 [0.187, 0.266] | .39 |
| Compassion | 0.121 [0.083, 0.175] | **.18** | 0.258 [0.079, 1.184] | .38 | 0.296 [0.248, 0.350] | .44 |
| Respectfulness | 0.049 [0.033, 0.072] | **.17** | 0.130 [0.036, 0.664] | .44 | 0.117 [0.094, 0.146] | .40 |
| Trust | 0.089 [0.059, 0.132] | **.22** | 0.109 [0.011. 0.697] | .27 | 0.211 [0.165, 0.263] | .52 |
| Conscientiousness | 0.206 [0.136, 0.304] | **.29** | 0.126 [0.006, 1.032] | .18 | 0.371 [0.303, 0.448] | .53 |
| Organization | 0.407 [0.298, 0.551] | **.38** | 0.093 [0.008, 0.506] | .09 | 0.581 [0.503, 0.671] | .54 |
| Productiveness | 0.122 [0.077, 0.184] | **.19** | 0.179 [0.048, 0.857] | .27 | 0.355 [0.290, 0.431] | .54 |
| Responsibility | 0.072 [0.047, 0.109] | **.14** | 0.254 [0.046, 1.540] | .51 | 0.175 [0.123, 0.230] | .35 |
| Negative emotionality | 0.024 [0.010, 0.044] | **.06** | 0.161 [0.038, 0.854] | .41 | 0.209 [0.162, 0.264] | .53 |
| Anxiety | 0.018 [0.001,0.043] | **.05** | 0.142 [0.019, 0.715] | .37 | 0.223 [0.168, 0.279] | .58 |
| Depression | 0.017 [0.010, 0.031] | **.09** | 0.083 [0.024, 0.379] | .43 | 0.095 [0.069, 0.123] | .49 |
| Emotional volatility | 0.024 [0.007, 0.047] | **.05** | 0.208 [0.041, 1.413] | .44 | 0.241 [0.190, 0.307] | .51 |
| Open-mindedness | 0.152 [0.106, 0.216] | **.36** | 0.057 [0.004, 0.408] | .13 | 0.219 [0.176, 0.273] | .51 |
| Aesthetic sensitivity | 0.172 [0.120, 0.245] | **.33** | 0.085 [0.016, 0.480] | .16 | 0.272 [0.216, 0.335] | .51 |
| Intellectual curiosity | 0.111 [0.071, 0.167] | **.28** | 0.100 [0.015, 0.528] | .25 | 0.182 [0.129, 0.231] | .46 |
| Creative imagination | 0.113 [0.079, 0.160] | **.32** | 0.035 [0.004, 0.238] | .10 | 0.202 [0.161, 0.248] | .58 |
| Narcissism | 0.234 [0.162, 0.331] | **.24** | 0.334 [0.097, 1.464] | .35 | 0.393 [0.322, 0.465] | .41 |
| Big Five trait mean | — | **.25** | — | .25 | — | .50 |
| Big Five facet mean | — | **.23** | — | .26 | — | .50 |

*Note.* $N$ = 1,139 observations (147 persons and their laptops by 8 independent observers). All variance components are significant, $p < .05$. CI = credibility interval. ICC = intraclass correlation coefficient. **Boldface** = consensus effect.

moderate mean of .23 (*SD* = .12). Narcissism consensus was small-to-moderate (.24). Although all variance components were technically significant, we caution that only ICCs ≥ .10 are practically meaningful. Thus, perceivers showed adequate consensus for all traits except negative emotionality and its three facets, which is consistent with the results above based on traditional methods.

*3.2.2 Accuracy*

We assessed accuracy by examining the latent association between two latent traits or facets—regressing target-level ratings of participants' personalities based on photos of their laptops for a given trait onto participants' self-reports of their own personality on the same trait (Table 4, left columns). Trait accuracy was moderate and significant for open-mindedness ($\beta$ = .35 [.15, .53]) and was small-to-moderate for the other four traits (range: .17 to .19) with a mean of $\beta$ = .21 (*SD* = .08). Facet accuracy was strongest for emotional volatility ($\beta$ = .38 [.05, .73]) and two open-mindedness facets—aesthetic sensitivity ($\beta$ = .36 [.17, .53] and intellectual curiosity ($\beta$ = .38 [.17, .57])—and was weakest for responsibility ($\beta$ = .01 [-.23, .25]), with a mean of $\beta$ = .18 (*SD* = .12).

We next ran the same accuracy models while controlling for Mac-ness and log number of stickers as covariates (Table 4, right columns). Mac-ness related (a) positively to latent ratings of extraversion (and its energy and sociability facets), agreeableness (and its facets), conscientiousness (and its facets), and narcissism; and (b) negatively to negative emotionality and its depression facet. Log number of stickers related (a) positively to extraversion (and its facets), agreeableness (and its trust and compassion facets), open-mindedness (and its facets), and narcissism; and (b) negatively to organization and depression. After controlling for both covariates,

**Table 4.** *CC-SEM Accuracy Slopes with (right) and without (left) Controlling for Laptop Mac-ness and Number of Stickers*

| | Accuracy | | Accuracy with Mac-ness and stickers covariates | |
|---|---|---|---|---|
| Personality trait/facet | Estimate | β | Estimate | β |
| Extraversion | 0.117 [-0.019, 0.261] | .18 [-.03, .37] | 0.117 [0.000, 0.238] | .17 [.00, .34] |
| Mac-ness | | | 0.344 [0.154, 0.537] | .28 [.12, .42] |
| Stickers | | | 0.503 [0.367, 0.642] | .57 [.43, .70] |
| Sociability | 0.074 [-0.032, 0.184] | .13 [-.06, .33] | 0.080 [-0.012, 0.173] | .14 [-.02, .31] |
| Mac-ness | | | 0.332 [0.144, 0.523] | .28 [.12, .42] |
| Stickers | | | 0.460 [0.322, 0.597] | .54 [.39, .66] |
| Assertiveness | 0.156 [0.036, 0.278] | .26 [.06, .44] | 0.152 [0.036, 0.269] | .25 [.06, .42] |
| Mac-ness | | | 0.100 [-0.120, 0.324] | .08 [-.10, .26] |
| Stickers | | | 0.250 [0.091, 0.407] | .28 [.10, .45] |
| Energy level | 0.073 [-0.074, 0.218] | .11 [-.11, .31] | 0.066 [-0.042, 0.180] | .10 [-.06, .25] |
| Mac-ness | | | 0.285 [0.142, 0.425] | .29 [.15, .43] |
| Stickers | | | 0.467 [0.362, 0.571] | .66 [.54, .77] |
| Agreeableness | 0.132 [-0.026, 0.293] | .19 [-.04, .39] | 0.101 [-0.052, 0.259] | .14 [-.07, .34] |
| Mac-ness | | | 0.244 [0.103, 0.386] | .32 [.14, .48] |
| Stickers | | | 0.112 [0.013, 0.212] | .21 [.02, .38] |
| Compassion | 0.031 [-0.096, 0.164] | .06 [-.17, .29] | 0.011 [-0.107, 0.138] | .02 [-.18, .24] |
| Mac-ness | | | 0.314 [0.150, 0.480] | .36 [.18, .53] |
| Stickers | | | 0.184 [0.069, 0.299] | .30 [.11, .47] |
| Respectfulness | 0.111 [-0.020, 0.246] | .20 [-.04, .42] | 0.097 [-0.032, 0.235] | .17 [-.06, .40] |
| Mac-ness | | | 0.158 [0.046, 0.275] | .29 [.08, .47] |
| Stickers | | | 0.055 [-0.019, 0.135] | .14 [-.05, .34] |
| Trust | 0.107 [-0.016, 0.226] | .23 [-.04, .45] | 0.099 [-0.014, 0.221] | .20 [-.03, .43] |
| Mac-ness | | | 0.230 [0.085, 0.379] | .31 [.12, .49] |
| Stickers | | | 0.127 [0.026, 0.227] | .24 [.05, .42] |
| Conscientiousness | 0.096 [-0.028, 0.220] | .17 [-.05, .38] | 0.076 [-0.044, 0.199] | .13 [-.08, .34] |
| Mac-ness | | | 0.398 [0.186, 0.618] | .35 [.17, .51] |
| Stickers | | | -0.018 [-0.170, 0.130] | -.02 [-.20, .16] |
| Organization | 0.094 [-0.023, 0.216] | .16 [-.04, .35] | 0.102 [-0.014, 0.211] | .16 [-.02, .34] |
| Mac-ness | | | 0.474 [0.195, 0.754] | .30 [.12, .46] |
| Stickers | | | -0.360 [-0.568, -0.169] | -.32 [-.48, -.15] |
| Productiveness | 0.061 [-0.057, 0.190] | .13 [-.12, .37] | 0.037 [-0.081, 0.151] | .07 [-.16, .30] |
| Mac-ness | | | 0.297 [0.126, 0.480] | .35 [.15, .53] |
| Stickers | | | 0.042 [-0.082, 0.164] | .07 [-.13, .26] |
| Responsibility | 0.003 [-0.106, 0.118] | .01 [-.23, .25] | 0.014 [-0.091, 0.124] | .03 [-.19, .26] |
| Mac-ness | | | 0.255 [0.118, 0.388] | .38 [.18, .55] |
| Stickers | | | 0.069 [-0.029, 0.166] | .15 [-.06, .34] |
| Negative emotionality | 0.033 [-0.022, 0.090] | .18 [-.12, .49] | 0.034 [-0.020, 0.091] | .18 [-.10, .49] |
| Mac-ness | | | -0.107 [-0.214, -0.005] | -.28 [-.55, -.01] |
| Stickers | | | -0.069 [-0.143, 0.006] | -.25 [-.52, .02] |
| Anxiety | 0.003 [-0.059, 0.071] | .02 [-.40, .40] | 0.006 [-0.057, 0.068] | .04 [-.32, .43] |
| Mac-ness | | | -0.027 [-0.147, 0.089] | -.08 [-.44, .26] |
| Stickers | | | -0.044 [-0.126, 0.035] | -.18 [-.54, .15] |
| Depression | 0.027 [-0.026, 0.081] | .13 [-.12, .37] | 0.027 [-0.021, 0.079] | .13 [-.10, .37] |
| Mac-ness | | | -0.109 [-0.196, -0.028] | -.33 [-.54, -.09] |
| Stickers | | | -0.101 [-0.157, -0.046] | -.43 [-.65, -.20] |
| Emotional volatility | 0.063 [0.009, 0.121] | .38 [.05, .73] | 0.057 [0.004, 0.111] | .33 [.02, .64] |
| Mac-ness | | | -0.045 [-0.165, 0.071] | -.11 [-.44, .18] |
| Stickers | | | 0.061 [-0.027, 0.142] | .22 [-.09, .52] |
| Open-mindedness | 0.281 [0.120, 0.444] | .35 [.15, .53] | 0.251 [0.107, 0.396] | .31 [.13, .47] |
| Mac-ness | | | 0.130 [-0.021, 0.284] | .14 [-.02, .28] |
| Stickers | | | 0.358 [0.247, 0.477] | .52 [.38, .65] |
| Aesthetic sensitivity | 0.222 [0.105, 0.336] | .36 [.17, .53] | 0.176 [0.064, 0.286] | .28 [.10, .44] |
| Mac-ness | | | 0.131 [-0.037, 0.304] | .13 [-.04, .29] |
| Stickers | | | 0.308 [0.187, 0.435] | .43 [.27, .57] |
| Intellectual curiosity | 0.247 [0.106, 0.394] | .38 [.17, .57] | 0.239 [0.104, 0.378] | .35 [.16, .54] |
| Mac-ness | | | 0.075 [-0.073, 0.230] | .09 [-.09, .27] |
| Stickers | | | 0.274 [0.160, 0.392] | .45 [.28, .61] |
| Creative imagination | 0.107 [-0.060, 0.271] | .15 [-.08, .36] | 0.148 [0.002, 0.303] | .20 [.00, .39] |
| Mac-ness | | | 0.126 [-0.010, 0.263] | .15 [-.01, .32] |
| Stickers | | | 0.357 [0.258, 0.456] | .61 [.46, .73] |
| Narcissism | 0.062 [-0.087, 0.220] | .09 [-.13, .31] | 0.065 [-0.062, 0.203] | .09 [-.09, .28] |
| Mac-ness | | | 0.273 [0.073, 0.486] | .23 [.07, .39] |
| Stickers | | | 0.460 [0.316, 0.601] | .54 [.38, .67] |

*Note.* *N*s = 1,139 observations of 147 participants and their laptops (targets) from 8 raters (perceivers). Brackets = 95% Bayesian credibility interval. Mac-ness = probability that laptop is an Apple Macintosh (vs. PC; 0, .25, .50, .75, 1). Stickers = natural log of number of laptop stickers. CC-SEM = cross-classified structural equation model. *$p < .05$.

significant accuracy effects emerged for both extraversion (β = .17 [.00, .34]) and creative imagination (β = .20 [.00, .39]); none of the latent bivariate accuracy effects became non-significant after controlling for covariates.

Only one self-reported trait or facet related to either covariate: Participants reporting higher aesthetic sensitivity decorated their laptops with significantly more stickers (*b* = 0.157 [0.011, 0.303]). This link allowed us to test statistical—but not necessarily causal (Bullock et al., 2010)—partial mediation because (a) the direct accuracy effect for aesthetic sensitivity was significant (*b* = 0.222 [0.105, 0.336] and (b) it remained significant after controlling for log stickers (*b* = 0.176 [0.064, 0.286], and log stickers related to ratings of aesthetic sensitivity (*b* = 0.308 [0.187, 0.435]. The indirect accuracy effect via log stickers was indeed significant (0.046 [0.003, 0.100] suggesting partial mediation (Figure 1). Thus, people with higher aesthetic sensitivity adorned their laptops with more stickers, and raters likely viewed laptops with more stickers as being an honest cue of their owners' aesthetic sensitivity.

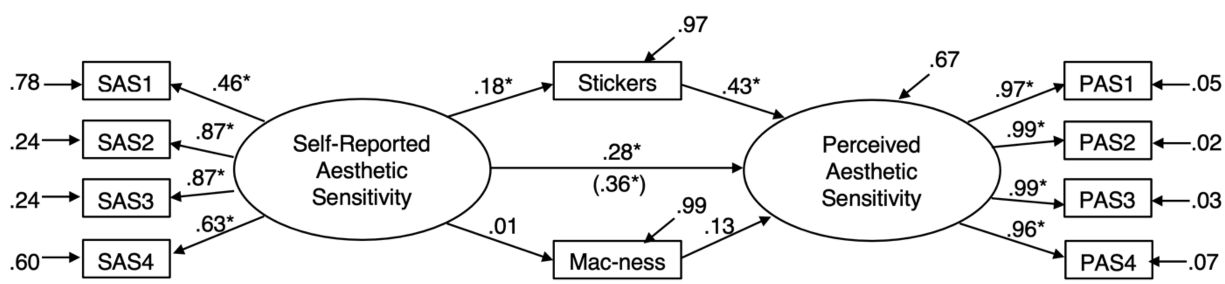


**Figure 1.** Results from the between-target level of a cross-classified structural equation model (CC-SEM) that examines log number of stickers as a cue for—or mediator of—the accuracy effect of the aesthetic sensitivity facet of open-mindedness. Standardized coefficients are shown; standardized coefficient of direct effect without cues or mediators appears in parentheses. Unstandardized coefficients and significant indirect effect via stickers are given in the main text. Based on 1,139 observations cross-classified across eight perceivers and 147 targets (laptops and their owners). SAS1–4: Self-reported aesthetic sensitivity items. PAS1–4: Perceived aesthetic sensitivity items. Mac-ness = probability that laptop is an Apple Macintosh (vs. PC; 0, .25, .50, .75, 1). Stickers = natural log of number of laptop stickers. **p* < .05.

## 4 Discussion

Raters unacquainted with laptop owners were able to accurately perceive their extraversion and open-mindedness based on their laptop stickers. But extraversion accuracy was significant only after controlling for laptop type and number of stickers, and driven largely by its assertiveness facet. In contrast, open-mindedness and its facets tended to have the strongest, most robust accuracy effect, with aesthetic sensitivity accuracy being partially statistically mediated by number of stickers as a possible cue. Unexpectedly, the accuracy effect for emotional volatility was significant despite showing poor consensus. Regarding consensus effects, with the exception of negative emotionality and its facets, most were in ranges consistent with prior studies of personality perception based on face-to-face zero-acquaintance, owned objects, or inhabited spaces. CC-SEMs yielded slightly smaller consensus—and slightly larger accuracy—correlations than traditional methods that ignore multiple dependencies and measurement error.

### *4.1 Limitations, Constraints on Generality, and Future Directions*

The present research has multiple potential limitations. First, despite having adequate power, having far more raters would have assured greater generalizability and allowed for testing for perceiver-level effects (e.g., age, gender, rater personality), which is fruitful avenue for future research. We chose to use eight raters to emulate Gosling et al.'s (2002) seminal studies on personality judgments of people's offices (8 raters) and bedrooms (7 raters) as well as Buffardi and Campbell's (2008) pioneering narcissism judgments based on Facebook profiles (5 raters). Greater numbers of raters allows for testing perceiver-level moderators, including gender and their own personality traits, which could be used to examine same- or cross-gender personality perception effects and personality projection or assumed similarity, where people overperceive their own traits in others (e.g., more agreeable people view others are similarly more agreeable; e.g., Kenny 2020; Webster & Campbell, 2021). In addition, more raters allows for greater generalizability of findings (see Judd et al., 2012; Wells & Windschitl, 1999).

A second limitation is that we did not collect additional sources of target personality such as peer reports, which could provide more definitive evidence of the validity of people's self-reported personality ratings (e.g., Naumann et al., 2009). Third, we did not set out to examine cues systematically (cf. Gosling et al., 2002). Nonetheless, that raters used number-of-stickers as a cue of laptop owners' aesthetic sensitivity (e.g., interest in art or literature) shows the potential explanatory power of assessing cues in this domain. Fourth, because we used laptops with three or more stickers, our findings may not generalize to those with fewer stickers. Fifth, our sample was homogenous by design. In the U.S., college students adorn their laptops with stickers; to this end, we chose undergraduates as both our targets and perceivers. Students likely choose and use laptop stickers to convey their interests specifically to their peers, rather than a broader demographic. Consequently, our findings should only be generalized to U.S. college students (Henrich et al., 2010; Simons et al., 2017).

Because our sample's gender imbalance (82% of targets were women), we chose not to examine gender differences in how targets' laptops—and hence their personality traits—were perceived. Prior studies have shown some meaningful gender differences in accuracy effects based on physical appearance in photographs, with perceivers' conscientiousness judgements being more accurate for male than female targets (e.g., Naumann et al., 2009). Future research involving laptops or similar personal accoutrements should aim to secure more balanced gender samples regarding both targets and perceivers. If broadcasting one's interests are indeed a proximal goal of sticker displays, and if attracting a mate with such displays is a more distal goal, then one might expect gender differences in both personality trait consensus and accuracy.

### *4.2 Implications for Theory and Methods*

Regarding theoretical implications, the present findings support ecological notions of social perception (McArthur & Baron, 1983), where accuracy in personality judgement is likely adaptive and necessary for trait differences to emerge, be socially relevant, and assist observers in predicting others' behaviors. That perceivers show acceptable consensus for most Big Five traits and significant accuracy for at least two traits—extraversion and open-mindedness—based solely on viewing people's laptop stickers is remarkable, but also understandable because laptop stickers likely serve as socially directed advertisements of their owners' preferences and identities. The present study also advances personality perception methods on three fronts: Ours is among the first studies to use (a) facets in addition to traits, (b) latent-variable models to account for measurement error, and (c) cross-classification analysis to optimally partition variance—or account for non-independent observations—at both the target and perceiver levels.

Regarding consensus methods, an advantage of the traditional approach is it requires only the calculation of the MIC or mean interrater (or interitem) correlation, which most program (and some R packages) include in their coefficient alpha or internal consistency reliability procedures. In contrast, an advantage of the variance-partitioning and ICC approach to assessing consensus using CC-SEM is that it accounts for multiple sources or levels of dependence (both targets and perceivers) and thus produces more accurate effects. But a key drawback of this approach is its computational complexity—it requires access to and familiarity with programs that can handle mixed, multilevel, or cross-classified data structures (e.g., lme4 [Bates et al., 2015], HLM [Raudenbush et al., 2019], Mplus [Muthén & Muthén, 2017]).

Regarding accuracy, an advantage of traditional methods is that both the aggregated-observer and single-observer approaches are computationally intuitive; the former simply requires taking the mean rating across perceivers before being correlated with targets' self-reports, whereas the latter examines individual-perceiver correlations with targets' self-reports and then averages those correlations together. A key drawback is that neither approach accounts for non-independence among raters or perceivers and needlessly throws away meaningful variance by collapsing or averaging across raters or perceivers. In contrast, because a CC-SEM or variance-partitioning approach explicitly models variance at the target and perceiver levels simultaneously, non-independence is accounted for and more precise, less-biased accuracy effects can be attained. Once again, a disadvantage of CC-SEMs are their computational complexity and their need for specialized packages or programs (e.g., Mplus). Moreover, much like any latent-variable modeling approach, multiple items, measures, or time points for each psychological construct are often required for CC-SEM model identification and convergence. Nevertheless, the increased accuracy and flexibility offered by CC-SEMs should encourage researchers who study personality judgement to consider their adoption so long as multiple assessments per construct are feasible (see Nestler & Back, 2017).

## Open Practices

Data, Mplus code, and variable code book are available at:

https://osf.io/4ygca/?view_only=47f4639812d3431282a444fd4ace5fd9

## References

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Biesanz, J. C., & Wallace, J. D. (2020). Accuracy in person perception. In P. J. Corr & G. Matthews (Eds.) *The Cambridge handbook of personality psychology* (2nd ed., pp. 44–55). New York: Cambridge University Press.

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

Buffardi, L. E., & Campbell, W. K. (2008). Narcissism and social networking web sites. *Personality and Social Psychology Bulletin, 34*(10), 1303–1314. https://doi.org/10.1177/0146167208320061

Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (Don't expect an easy answer). *Journal of Personality and Social Psychology, 98,* 550–558. https://doi.org/10.1037/a0018933

Claus, A. M., Arend, M. G., Burk, C. L., Kiefer, C., & Wiese, B. S. (2020). Cross-classified models in I/O psychology. *Journal of Vocational Behavior*, 103447. https://doi.org/10.1016/j.jvb.2020.103447

Funder, D. C. (1995). On the accuracy of personality judgement: A realistic approach. *Psychological Review, 102*(4), 652–670. https://doi.org/10.1037/0033-295X.102.4.652

Galton, F. (1884). Measurement of character. *The Fortnightly Review, 36,* 179–185.

Gillath, O., Bahns, A. J., Ge, F., & Crandall, C. S. (2012). Shoes as a source of first impressions. *Journal of Research in Personality*, *46*(4), 423–430. https://doi.org/10.1016/j.jrp.2012.04.003

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*(3), 379–398. https://doi.org/10.1037/0022-3514.82.3.379

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

Hall, J. A., Mast, M. S., & West, T. (2016). Accurate interpersonal perception: Many traditions, one topic. In J. A. Hall, M. S. Mast, and T. West (Eds.), *The social psychology of perceiving others accurately* (pp. 3–22). Cambridge, UK: Cambridge University Press.

Harari, G. M., Graham, L. T., & Gosling, S. D. (2015). Personality impressions of world of Warcraft players based on their avatars and usernames: Consensus but no accuracy. *International Journal of Gaming and Computer-Mediated Simulations*, *7*(1), 58–73. http://doi.org/10.4018/IJGCMS.2015010104

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*(2–3), 61–83. http://doi.org/10.1017/S0140525X0999152X

John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–138). New York: Guilford.

Jonason, P. K., & Webster, G. D. (2010). The Dirty Dozen: A concise measure of the Dark Triad. *Psychological Assessment, 22*(2), 420–432. https://doi.org/10.1037/a0019265

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103,* 54–69. http://doi.org/10.1037/a0028347

Kenny, D. A. (2020). *Interpersonal Perception* (2nd ed.). New York: Guilford.

McCain, J. L., & Campbell, W. K. (2018). Narcissism and social media use: A meta-analytic review. *Psychology of Popular Media Culture*, *7*(3), 308–327. https://doi.org/10.1037/ppm0000137

Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Personality judgments based on physical appearance. *Personality and Social Psychology Bulletin, 35*(12), 1661–1671. https://doi.org/10.1177/0146167209346309

Nave, G., Minxha, J., Greenberg, D. M., Kosinski, M., Stillwell, D., & Rentfrow, J. (2018). Musical preferences predict personality: Evidence from active listening and Facebook likes. *Psychological Science, 29*(7), 1145–1158. https://doi.org/10.1177/0956797618761659

Nestler, S., & Back, M. D. (2013). Applications and extensions of the lens model to understand interpersonal judgments at zero acquaintance. *Current Directions in Psychological Science, 22*(5), 374–379. https://doi.org/10.1177/0963721413486148

Nestler, S., & Back, M. D. (2017). Using cross-classified structural equation models to examine the accuracy of personality judgments. *Psychometrika, 82*(2), 475–497. https://doi.org/10.1007/s11336-015-9485-6

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R. T., Jr., & du Toit, M. (2019). *HLM8: Hierarchical linear and nonlinear modeling*. Skokie, IL: Scientific Software International.

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science, 12*(6), 1123–1128. https://doi.org/10.1177/1745691617708630

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology, 113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Tskhay, K. O., & Rule, N. O. (2014). Perceptions of personality in text-based media and OSN: A meta-analysis. *Journal of Research in Personality, 49,* 25–30. https://doi.org/10.1016/j.jrp.2013.12.004

Webster, G. D., & Campbell, J. T. (2021). Personality perception in *Game of Thrones*: Character consensus and personality projection. Invited resubmission, *Psychology of Popular Media*. https://doi.org/10.31234/osf.io/8xnd5

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin, 25,* 1115–1125. https://doi.org/10.1177/01461672992512005