Journal of Personality 82:5, October 2014

DOI: 10.1111/jopy.12075

An Examination of Information Quality as a Moderator of Accurate Personality Judgment

Tera D. Letzring[1] and Lauren J. Human[2]
[1]Idaho State University
[2]University of British Columbia

Abstract

Information quality is an important moderator of the accuracy of personality judgment, and this article describes research focusing on how specific kinds of information are related to accuracy. In this study, 228 participants (159 female, 69 male; mean age = 23.43; 86.4% Caucasian) in unacquainted dyads were assigned to discuss thoughts and feelings, discuss behaviors, or engage in behaviors. Interactions lasted 25–30 min, and participants provided ratings of their partners and themselves following the interaction on the Big Five traits, ego-control, and ego-resiliency. Next, the amount of different types of information made available by each participant was objectively coded. The accuracy criterion, composed of self- and acquaintance ratings, was used to assess distinctive and normative accuracy using the Social Accuracy Model. Participants in the discussion conditions achieved higher distinctive accuracy than participants who engaged in behaviors, but normative accuracy did not differ across conditions. Information about specific behaviors and general behaviors were among the most consistent predictors of higher distinctive accuracy. Normative accuracy was more likely to decrease than increase when higher-quality information was available. Verbal information about behaviors is the most useful for learning about how people are unique.

You can discover more about a person in an hour of play than in a year of conversation.—Plato

Was Plato correct—is information based on play really more useful for learning about a person than information based on conversation? What is the best way to learn about someone, and what kind of information is most useful for this endeavor? The answers to these questions are important because people often make decisions about others—such as whom to hire, marry, or allow to watch one's children—and these decisions are likely to be based in part on judgments of personality. Accurate judgments of personality should result in better decisions. People should know about the kinds of information they should seek in order to improve the accuracy of their judgments, as well as when to reserve judgment until they are able to obtain more of the right types of information. However, little is known about the specific kinds of information that are most useful for making accurate personality judgments.

Personality psychologists have a wealth of information about the types of cues and behaviors that are related to personality traits based on psychological assessment instruments. This knowledge falls under the first process in Brunswik's (1956) lens model, in which cues about the target of judgment are made available. However, most people do not have access to the results of psychological tests when making judgments of personality and therefore do not know which cues are ecologically valid, or related to the personality trait being judged. For this reason, it is important to learn how laypeople make accurate personality judgments. Previous studies using the framework of the lens model have examined how behaviors serve as personality cues, finding evidence for the validity of a variety of behavioral cues as indicators of personality (Funder & Sneed, 1993; Gangestad, Simpson, DiGeronimo, & Biek, 1992; Gifford, 1994).

Another useful model for understanding how accuracy of personality judgment is achieved is Funder's Realistic Accuracy Model (RAM; Funder, 1995, 1999). Funder proposed four moderators of accuracy: good judge, target, trait, and information. He specified two aspects of good information: quantity and quality. There is consistent support for the importance of information quantity for accurate judgment and the acquaintanceship effect (Biesanz, West, & Millevoi, 2007; Colvin & Funder, 1991; Letzring, Wells, & Funder, 2006). Earlier research did not decompose accuracy into components, but

This research was supported by Grant No. 1005 from the Faculty Research Committee, Idaho State University, Pocatello, Idaho.

Correspondence concerning this article should be addressed to Tera D. Letzring, Department of Psychology, Idaho State University, 921 S. 8th Ave., Pocatello, ID 83209. Email: letztera@isu.edu.

more recently it was found that distinctive accuracy (understanding targets' unique ordering of traits and how they are different from other targets and the average person) increases with greater information, whereas normative accuracy (perceiving targets in a way that is similar to what the average person is like) decreases or stays the same with greater information (Biesanz & Human, 2010; Biesanz et al., 2007).

Information quality, on the other hand, has been the focus of less research attention. Based on RAM, information is high quality, or useful for making accurate personality judgments, when it is *available* to the judge and *relevant* to the trait being judged, thereby facilitating the first two stages of accurate person perception. There is evidence that information from both unstructured and getting-to-know-you conversations results in greater accuracy than information from responding to trivia questions, suggesting that higher-quality information was available in the first two types of conversations (Letzring et al., 2006). Additionally, Beer and Brooks (2011) found that information about values generated higher accuracy for Neuroticism; information about facts that made the target unusual generated higher accuracy for Conscientiousness; and information about values and facts generated about equal levels of accuracy for Extraversion, Openness, and Agreeableness. Overall, certain types of information that are made available through conversation appear to be relevant to personality traits. However, many questions about what constitutes high-quality information remain, making it critical to more systematically investigate what types of information are of high quality.

One aspect of information quality to consider is what types of information are the most available to others, and therefore potentially of high quality. Directly observable behaviors and characteristics are the most available types of information, making these good candidates for high-quality information. This may explain why more observable traits (such as Extraversion) are generally judged more accurately than less observable traits (such as Neuroticism; Funder & Dobroth, 1987; Human & Biesanz, 2011a; John & Robins, 1993). Research has examined the usefulness of various types of directly observable information for making accurate judgments of personality (Borkenau & Liebler, 1995; Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004; Naumann, Vazire, Rentfrow, & Gosling, 2009).There is ample evidence that highly available behavioral and physical cues contribute to forming accurate impressions of some traits, suggesting that this type of information is high quality for at least some traits.

However, there are other traits for which there are fewer cues that are directly observable. And in many cases, first impressions are based on conversations with new acquaintances, in addition to observations of behavior or other physical cues. Conversations may increase the availability of information that is typically less directly observable, such as internal thoughts and feelings. To the extent that thoughts and feelings provide relevant information, the provision of such information may be particularly useful in aiding accuracy of impressions. In contrast, discussing more directly observable information, such as behaviors, may be less useful for improving accuracy beyond simply observing someone's behavior. Indeed, Andersen (1984) found that self-other agreement (a commonly used measure of accuracy) was higher among judges exposed, via audiotaped interviews, to information about thoughts and feelings of the target than for judges exposed to information about hobbies and activities. Based on this logic, discussions of thoughts and feelings may be especially useful for judging traits that have fewer directly observable behaviors related to them (such as Neuroticism) because the judge is less likely to have access to many high-quality cues from behavioral observation alone. On the other hand, discussion of thoughts and feelings may be less useful for making accurate judgments of more behaviorally oriented traits, and instead, discussions about behaviors or actually engaging in behavior may be more useful for accurately judging these traits.

The Present Research

The goal of the current project was to systematically compare the usefulness of different types of information for making accurate personality judgments, including verbal information about thoughts and feelings, verbal information about behaviors, and behavioral information based on engaging in behaviors. Participants interacted in dyads and were randomly assigned to one of four experimental conditions in an attempt to manipulate the type of information that was available to their partner. Following the interactions, participants provided ratings of each other and self-ratings of personality. After all interactions had occurred, the amounts of different types of information that were made available by each participant were coded.

The criteria for accuracy consisted of a composite of self- and acquaintance ratings. Two types of accuracy were examined: distinctive accuracy and normative accuracy (Biesanz, 2010; Furr, 2008), which are analogous to Cronbach's (1955) components of differential and stereotype accuracy, respectively. Distinctive accuracy refers to understanding a specific target's unique ordering of traits (e.g., he is more talkative than kind) as well as how he or she is different from other specific targets and the average person (e.g., she is more talkative than the average person). Normative accuracy refers to perceiving the target in a way that is similar to what the average person is like. Normative accuracy is based in part on the judges' expectations about what the average person is like (Rogers & Biesanz, 2013). Because people on average are by definition similar to the average person, viewing a target in this manner is likely to enhance overall accuracy, especially in first impressions contexts when information is limited. Note, however, that the normative profile is also very positive in nature (e.g., Edwards, 1957), and therefore normative accuracy may also reflect viewing the target more positively.

Hypotheses

Three competing hypotheses were tested. Hypothesis 1 was that as the availability of verbal information about thoughts and feelings and verbal information about behaviors increased, so too would distinctive accuracy. Additionally, it was predicted that conditions in which participants were instructed to have a conversation would be associated with higher distinctive accuracy than conditions in which participants were instructed to engage in behaviors because it is likely that discussions will make more information available about thoughts, feelings, and behavior. However, increases in distinctive accuracy are likely to coincide with decreases in normative accuracy because judgments would rely less on knowledge of what the average person is like as more ecologically valid information about the individual became available (e.g., Biesanz et al., 2007). Therefore, it is predicted that more information about thoughts, feelings, and behaviors would be related to lower normative accuracy. This hypothesis is in line with Funder's RAM because increasing the success of the relevance and availability stages by increasing information about thoughts, feelings, and behaviors should increase distinctive accuracy and decrease normative accuracy.

Hypothesis 2 was that information about thoughts and feelings would be the most useful for achieving distinctive accuracy. This prediction was based on the previous finding that listening to information about thoughts and feelings resulted in higher self-other agreement than listening to information about hobbies and activities (Andersen, 1984). As in Hypothesis 1, the prediction for normative accuracy was that it would be negatively related to the availability of information about thoughts and feelings. A more focused version of this hypothesis is that traits that are more internal and less observable, such as Neuroticism and Openness, are especially likely to be judged with more distinctive accuracy when more information about thoughts and feelings is available because these traits have fewer cues related to behavior. As such, this hypothesis may not hold for highly observable traits, such as Extraversion, for which behavioral information may be the most beneficial (see Hypothesis 3). Hypothesis 2 differs from Hypothesis 1 in that it includes the prediction that information about thoughts and feelings is more useful for distinctive accuracy than is information about behaviors, whether that is accomplished through discussion or behavioral observation.

Hypothesis 3, based on the opening quote by Plato, was that participants who engaged in behaviors would be seen with the highest levels of distinctive accuracy and the lowest levels of normative accuracy. Thus, this hypothesis stands in direct opposition to Hypothesis 1. A more focused version of this hypothesis is that traits that have more behavioral cues and are more observable, such as Extraversion, are especially likely to be judged with more distinctive accuracy when behavior is directly observed because this information may be the most relevant to these types of traits. Evidence confirming this hypothesis would suggest that direct behavioral observation results in higher-quality information than discussions about either behavior or thoughts and feelings.

These hypotheses were tested in two ways: with an experimental manipulation of information availability and based on objective codings of available information. The current study extends previous research by examining accuracy following an interaction between pairs of judges and targets (as compared to having judges listen to taped interviews of targets) and by including a condition in which participants engage in a series of behaviors (rather than the focus being on discussion). Furthermore, the types of information that were actually provided in the interactions were objectively coded by research assistants, which allowed us to examine the success of the manipulation as well as to examine more precisely how the amounts of different types of information were predictive of accuracy. Finally, we examined both distinctive and normative accuracy, which is important because the predictions differ for the type of accuracy being considered.

METHOD

Participants

Participants were 228 students (159 female, 69 male) at Idaho State University who participated in exchange for credit in a psychology class.[1] The average age of participants was 23.43 years ($SD = 6.72$). Participants were primarily of Caucasian ethnicity (86.4%; 3.9% Hispanic, 1.8% Asian, and 7.9% other).

Additionally, 308 acquaintance-informants provided information about one of the participants. One participant had four acquaintance-informants, 10 participants had three acquaintance-informants, 101 participants had two acquaintance-informants, 71 participants had one acquaintance-informant, and 45 participants had zero acquaintance-informants. Participants with zero acquaintance-informants were excluded from the analyses because it was not possible to compute realistic accuracy without more than just the self-ratings from these participants. This resulted in a total of 183 particpants.

Finally, research assistants were trained to provide objective codings of the videotaped interactions. Over five semesters, 14 coders (six females, eight males) completed between 21 and 194 codings each.

Measures

The International Personality Item Pool. The International Personality Item Pool 300-item version of the NEO PI-R facets (IPIP NEO PI-R facets; International Personality Item Pool, n.d.). Self-report and other-report formats of this questionnaire were used. The IPIP facet scores correlate with the NEO PI-R facet sores within the range of .60 (dutifulness facet of Conscientiousness) and .81 (assertiveness facet of Extraversion),

and the alpha reliabilities of the facet scores based on the IPIP range from .71 (activity level facet of Extraversion and dutifulness facet of Conscientiousness) to .88 (anger and depression facets of Neuroticism; International Personality Item Pool, n.d.).

Ego-Control and Ego-Resiliency. The 51-item Ego-Control and Ego-Resiliency scale (ECER; Letzring, Block, & Funder, 2005) assesses ego-control, or one's level of impulse control, and ego-resiliency, or how much a person adapts the level of control to fit the situation. The scale has sufficient internal reliability for ego-control ($\alpha = .63$) and ego-resiliency ($\alpha = .72$; Letzring et al., 2005), and high test-retest reliability for ego-resiliency over a 5-year time span, $r = .67$ and .51 for females and males, respectively (Block & Kremen, 1996). Reliability in the current study was adequate for both ego-control ($\alpha = .82$) and ego-resiliency ($\alpha = .71$).[2]

Procedures

Manipulation of Type of Information. Dyads were randomly assigned to one of four conditions in an attempt to manipulate the type of information that would be available during the interaction. Participants in the Thoughts and Feelings condition were instructed to talk about their thoughts and feelings in several situations (i.e., with family, with friends, while at school or work, while making a difficult decision, while at social events, when in a good mood, when in a bad mood, and yesterday). Participants in the Behavior Discussion 1 condition were instructed to talk about the behaviors they typically engage in while in the same situations as the Thoughts and Feelings condition. The Thoughts and Feelings and Behavior Discussion 1 conditions were designed so that the same situations were discussed, but in the first condition thoughts and feelings were the focus and in the second condition behaviors were the focus. Participants in the Behaviors condition were instructed to participate in several behaviors, such as a cooperative interaction, reading a poem out loud, telling stories based on cards from the Thematic Apperception Test (TAT; Murray, 1943), explaining the meaning of sayings from a list (such as "People who live in glass houses should not throw stones" and "Those who criticize our generation forget who raised it"), talking about whatever they would like for 3 min, and playing a competitive game in which players take turns removing blocks from a stack until the stack falls. Participants in the Behavior Discussion 2 condition described their typical behaviors in several situations (i.e., when cooperating, when giving a presentation or talking in front of people, a time when they were creative, when with friends, and when competing). The Behaviors and Behavior Discussion 2 conditions were designed so that similar situations were included, but in the Behaviors conditions the behaviors were performed whereas in the Behavior Discussion 2 condition the typical behaviors in these situations were discussed. This manipulation is especially useful for seeing how instructions about the kinds of information to discuss influence accuracy.

Data were collected for 28 dyads in the Thoughts and Feelings, Behavior Discussion 2, and Behaviors conditions, and for 30 dyads in the Behavior Discussion 1 condition.

Judgment Task and Self-Ratings. Following all types of interactions, participants moved to separate computers where they could not see each other and engaged in the judgment task. They first rated the personality of their partner using other-report versions of the IPIP NEO PI-R facets scale and the ECER scale. Next, they completed self-reports of these same measures, followed by demographic questions. Finally, they provided names and contact information for two family members and two close friends and were told that the researcher would contact these people to ask them to provide online personality ratings of the participant.

Coding of the Interactions. The main purpose of the coding was to assess the amount of information about thoughts and feelings or behaviors (both specific and general behaviors) that was available for each participant. Each item was coded on a 4-point scale ranging from 1 (*none*) to 4 (*a lot*). Several other types of information were also coded, including whether the participant introduced himself or herself (coded yes or no), information about other people the participant knows, information about relationships the participant has, information about events currently happening in the participant's life, information about past events in the participant's life, statements about the participant's own personality traits, expressiveness of the participant (in terms of facial expression, tone of voice, gestures, etc.), how personal the information was, how many of the assigned topics were discussed, how many additional topics were discussed,[3] how similar the participants were to each other (in terms of interests, reactions to events, personality, etc.), how much the participant talked about being similar to his or her partner, amount of eye contact, how quickly the participant behaved, and the total time of the interaction. Coders watched each interaction while focusing on only one participant. They could take notes during the interaction and pause or rewind the interaction if needed. Following the observation, they completed all codings for the participant on whom they had focused. If both participants in the same dyad were coded by the same person, the coder watched the interaction again while focusing on the other participant. Five participants were coded by three coders, 206 were coded by four coders, and 15 were coded by five coders.

Reliabilities for codings were computed for participants with four or five coders and based on only the first four coders for participants with five coders. The reliabilities were adequate for the amount of information about thoughts and feelings and general behaviors (Cronbach's alphas = .70 and .68, respectively), but low for specific behaviors (Cronbach's alpha = .44; see Table 1). Reliabilities were also adequate for information about other people, relationships, past events, the

Table 1 Descriptive Statistics for Interaction Coding

Item	Mean	*SD*	Alpha Reliability
Thoughts and feelings	2.87	.77	.70
Specific behaviors	1.88	.55	.44
General behaviors	2.70	.78	.68
Other people	2.31	.74	.77
Relationships	1.93	.64	.72
Current events	2.38	.61	.60
Past events	2.35	.70	.72
Own personality	2.11	.76	.77
Expressivenesss	2.57	.68	.72
Personal-ness of information	1.59	.54	.73
Assigned topics	6.44	1.28	.89
Additional topics	5.59	3.22	.64
Similarity of participants	1.68	.57	.70
Talk about similarity	1.68	.57	.62
Eye contact	3.02	.69	.74
Tempo/pace	2.83	.59	.72

Note. SD = standard deviation.

participant's own personality, expressiveness, personal-ness of information, number of assigned topics discussed, similarity of the participants, amount of eye contact, and the tempo/pace of behavior (Cronbach's alphas = .70 to .89); but they were somewhat low for information about current events, number of additional topics discussed, and amount of talk about similarity of participants to each other (Cronbach's alphas = .60 to .64).

Realistic Accuracy Criteria

Acquaintance-informants were contacted by email or telephone by a trained research assistant. They were informed about the nature of the study and told the name of the person who identified them as a family member or close friend. They were told that it should take about 30 minutes to complete the online questionnaires and that they would be paid $5.00 for their time after they completed the questionnaires. If they agreed to participate, they were given information about how to access the questionnaires on SurveyMonkey.com and asked to complete the questionnaires within the next 2 weeks. Acquaintance-informants who had not completed the questionnaires within 2 weeks were contacted again and reminded to participate. The questionnaires consisted of the IPIP NEO PI-R facets scale and the ECER scale in other-report formats.

When two or more informants responded, ratings were first averaged and then this average was averaged with the self-rating to form the realistic accuracy criteria.[4] Using this composite rating to assess accuracy is advantageous to using only a single rating (such as a self-report) because the composite rating is more likely to reflect what the target is actually like than is a single rating (Funder, 1995, 1999; Letzring et al., 2006). Items from the 300-item measure (which was used by targets and acquaintances) and the 50-item measure (which was used by judges) were compared to determine overlapping items, which resulted in 35 total items that were used to compute realistic accuracy (five items for Extraversion, six for Agreeableness, six for Conscientiousness, 10 for Neuroticism, and eight for Openness).

Analyses

Participants who were missing more than 20% of the items for the self-rating or the acquaintance-ratings were excluded from the analyses. Therefore, the final analyses were conducted for 167 participants for the Big Five traits and 159 participants for ego-control and ego-resiliency.

The data were analyzed following the social accuracy model procedures (SAM; see Biesanz, 2010, for a detailed explanation of the model; see Human & Biesanz, 2011b, for a detailed empirical example). This model is based on multilevel modeling and can be used to examine distinctive and normative accuracy within the same model. The data used for the analyses were at the item level, which allows for an examination of accuracy for separate traits because there are multiple items per trait. In the basic model, judges' ratings of the personality of the interaction partner were simultaneously predicted by (a) the accuracy criteria for each target on each item (reflecting distinctive accuracy), after subtracting out the average of all participants' self-reports on that item (i.e., grand-mean centering each item), and (b) by the average of all participants' self-reports ($N = 233$) on each item (reflecting normative accuracy). Items were not reverse-coded prior to analysis. In order to account for the potential dependence among the dyadic pairs who interacted with each other, distinctive and normative accuracy slopes were allowed to vary randomly across dyads as follows:

$$Y_{jk} = \beta_{0j} + \beta_{1j} TVal_{jk} + \beta_{2j} Mean_k + \varepsilon_{jk} \tag{1.1}$$

$$\begin{aligned} \beta_{0j} &= \beta_{00} + \beta_{01} Cond_j + u_{0j} \\ \beta_{1j} &= \beta_{10} + \beta_{11} Cond_j + u_{1j} \\ \beta_{2j} &= \beta_{20} + \beta_{21} Cond_j + u_{2j} \end{aligned} \tag{1.2}$$

In this model, Y_{jk} corresponds to each dyad j member's rating of his or her interaction partner on item k. $TVal_{jk}$ represents each dyad j member's personality accuracy criterion on item k. The intercept is represented by β_{0j}. The regression coefficient for the distinctive accuracy slope is β_{1j}, or the relationship between each member of dyad j's standing on item k predicting his or her interaction partner's rating of him or her on the same item k, after partialing out the normative profile (the mean target self-report for item k; $Mean_k$). β_{2j} is the regression coefficient for the normative accuracy slope, or the relationship between the mean participant self-report for item k predicting each dyad i member's rating of the same item k.

To compare distinctive and normative accuracy across the experimental conditions, we took a group codes approach

Table 2 Overall Levels of Accuracy and Effects of Experimental Conditions

	Overall Accuracy Levels (*b*)		Behavior Disc. 1 vs. Thoughts and Feelings (Gelman's *d*)		Behavior Disc. 2 vs. Thoughts and Feelings (Gelman's *d*)		Behaviors vs. Thoughts and Feelings (Gelman's *d*)		Behavior Disc. 1 vs. Behaviors (Gelman's *d*)		Behavior Disc. 2 vs. Behaviors (Gelman's *d*)	
	Dist.	Norm.	Dist.	Norm.	Dist.	Norm.	Dist.	Norm.	Dist.	Norm.	Dist.	Norm.
All Big Five traits	.18***	.88***	.17	–.07	.01	–.14	–.74***	–.17	.91**	.11	.75**	.03
Extraversion	.25***	.79***	.61	–.01	–.04	.14	–1.44**	–.16	2.05***	.15	1.40**	.30
Agreeableness	.14***	.95***	.09	–.15	–.27	–.35	–.63	–.14	.72	.01	.36	–.21
Conscientiousness	.12***	.92***	.78	–.08	.95	–.18	–.07	–.15	.85	.07	1.02	–.03
Neuroticism	.21***	.98***	–.16	.21	–.29	.06	–.60	.09	.44	.11	.31	.09
Openness	.38***	.49***	.12	.15	.29	.17	–.29	–.31	.41	.46	.58	.48
Ego-control	.24***	.83***	–.11	–.38	.18	–1.15**	–.95**	–.42	.84*	.04	1.14***	–.73
Ego-resiliency	.19***	1.12***	.02	–.95	–.30	–2.75	–1.05	–3.65	1.07	2.70	.75	.90

Note. $N = 181$. Disc. = discussion; Dist. = distinctive accuracy; Norm. = normative accuracy. *b* = unstandardized regression coefficients. Gelman's *d* = standardized regression coefficients, calculated as the change in the respective slope between the reference condition and the comparison condition, divided by the dyadic distinctive or normative accuracy random effect estimate standard deviation, as applicable. Note that the sign of the experimental effects are positive if the top row condition is higher and negative if the bottom row condition is higher.
***$p < .001$. **$p < .01$. *$p < .05$.

using the Thoughts and Feelings condition as the reference group. Specifically, we created three dummy-coded variables to indicate whether the participant was in the Behavior Discussion 1 condition (BD1: 0 = no, 1 = yes), Behavior Discussion 2 condition (BD2: 0 = no, 1 = yes), or Behaviors condition (B: 0 = no, 1 = yes). Each of these variables was then included as moderators of normative and distinctive accuracy within the same analysis. Thus, each interaction coefficient reflects whether accuracy was different in the relevant condition compared with the Thoughts and Feelings condition. Within the equations, $Cond_j$ represents dyad *j*'s condition (e.g., BD1). In the equation for β_{1j}, β_{11} represents the regression coefficient for condition moderating distinctive accuracy. Likewise, in the equation for β_{2j}, β_{21} represents the regression coefficient for condition moderating normative accuracy. Here, negative values for β_{11} and for β_{21}, when $Cond_j$ refers to BD1 (0 = Thoughts and Feelings condition; 1 = Behavior Discussion 1 condition), would suggest lower distinctive and normative accuracy, respectively, for dyads in the Behavior Discussion 1 condition relative to the Thoughts and Feelings condition. In both equations, u_{1j} and u_{2j} represent the dyadic random effects for distinctive and normative accuracy main effects, respectively, averaged across dyads. Note that these dyadic random effects include perceiver and target effects because they cannot be estimated separately given that the data are composed of single dyadic pairs (i.e., each participant interacted with only one other participant).

Table 2 presents unstandardized parameter estimates, *b*s, for mean levels of distinctive and normative accuracy (β_{1j} and β_{2j}, respectively) and standardized parameter estimates (Gelman's *d*s)[5] for all moderator effects, calculated by dividing the unstandardized regression coefficient by the dyadic distinctive or normative accuracy random effect estimate (u_{1j} and u_{2j}, respectively) standard deviation. Inferences were based upon the asymptotic *z*-test, and z-scores were obtained by dividing parameter estimates by normal theory standard errors.

Next, the relationships between the amounts of available information and distinctive and normative accuracy were assessed. All information indicators from the objective codings were standardized prior to the analyses and then included in the multilevel models as moderators of distinctive and normative accuracy, parallel to the procedure for examining the effects of condition. That is, instead of including condition as a moderator of distinctive and normative accuracy, the continuous information ratings were included as moderators. A positive, significant interaction would indicate that the corresponding information indicator is associated with significantly greater levels of distinctive or normative accuracy.

Tables 3 and 4 present effect size estimates, Gelman's *d*s, for these moderated effects, calculated as the change in the respective accuracy slope for a two standard deviation change in the information indicators divided by the respective random effect estimate standard deviation. This approach helps to make the effect size estimates for continuous variables comparable to effect size estimates for dichotomous predictors (e.g., Cohen's *d*), enabling a cleaner comparison between the effect size estimate for these continuous ratings and the effects of condition assignment (see Gelman, 2008; Human & Biesanz, 2011a, 2011b).

RESULTS

Manipulation Check

The amount of information about thoughts and feelings, specific behaviors, and general behaviors that was available in each condition (based on the objective codings) was compared to determine how well participants followed the instructions about what to discuss. For each dyad, the average of each type of information was calculated, and this number was used as the dependent variable in a one-way analysis of variance (ANOVA), with condition type as the independent variable.

Table 3 Type of Information as Moderator of Distinctive Accuracy

	Trait Being Judged							
Type of Information	All Big Five	E	A	C	N	O	EC	ER
Thoughts and feelings	.06	.47	.41	.49	–.01	.51**	.59**	.48
Specific behaviors	.55***	.75*	.12	.53	.41*	.50**	.35	.12
General behaviors	.43**	.64*	.43	.19	.20	.39*	.49*	.55
Other people	–.14	.44	–.12	.40	–.19	.14	.31	.14
Relationships	.24	.44	.44*	.29	.27	.08	.30	.04
Current events	.86***	.46	.35	–.24	.69***	.40*	.44*	.15
Past events	.36**	–.18	.44*	.61	.20	.44**	.07	–.34
Own personality	.71***	1.36***	.54*	.97**	.48**	.23	.60***	.60
Expressivenesss	–.37**	–.47	–.32	–.71*	–.31*	–.10	–.13	–.04
Personal-ness	.47***	.49	.32	.84**	.30	.21	.22	.26
Assigned topics	.02	.76*	.20	.09	.13	–.02	.11	.58
Additional topics	–.19	.16	.19	–.66	–.64**	.83	.35	–.21
Participant similarity	–.23	–.14	.05	.34	–.83	.34	.20	.10
Talk about similarity	–.02	.27	.22	.46	.08	.20	.67***	–.25
Eye contact	–.41**	–.09	–.36	–.11	–.38*	.26	.65**	.17
Tempo/pace	–.31*	–.64*	–.37	–.30	.00	–.28	–.23	–.73

Note. $N = 181$. E = Extraversion; A = Agreeableness; C = Conscientiousness; N = Neuroticism; O = Openness; EC = ego-control; ER = ego-resiliency. Standardized regression coefficients, Gelman's *d*, are reported, calculated as the change in the distinctive accuracy slope for a two standard deviation change in the information indicators, divided by the dyadic distinctive accuracy random effect estimate standard deviation, in order to make these comparable to the effect size estimates comparing conditions (Gelman, 2008).
***$p < .001$. **$p < .01$. *$p < .05$ $^{+} < .10$.

Table 4 Type of Information as Moderator of Normative Accuracy

	Trait Being Judged							
Type of Information	All Big Five	E	A	C	N	O	EC	ER
Thoughts and feelings	.37***	.61***	–.09	.13*	.36*	.62**	.31	2.79*
Specific behaviors	–.29***	.25*	–.46***	–.10*	.01	.38	.53	.44
General behaviors	–.29***	.32*	–.57***	–.15**	–.10	.36	–.10	.59
Other people	–.04	.35**	–.46***	–.07	.34*	.17	–.03	1.05
Relationships	–.28***	.33**	–.51***	–.21***	–.17	.20	.43	2.32
Current events	–.14	.02	–.10	–.05	–.10	.43*	.23	.02
Past events	.19**	.35**	–.10	.05	.42**	.37	–.09	–2.86*
Own personality	–.23**	.13	–.73***	–.02	–.10	.18	.25	2.67*
Expressivenesss	–.38***	.24*	–.51**	–.28***	–.12	–.11	–.16	1.01
Personal-ness	–.13*	.15	–.27**	–.09*	–.09	.13	.40	–.34
Assigned topics	.68***	.33	.11	.25**	.36	.10	.70	2.72*
Additional topics	.73***	.54***	.03	.14*	.54**	.65**	–.69	–.92
Similarity of participants	.10	.36*	–.30	.11	.35	.45	–.25	.60
Talk about similarity	.30	.14	.17	.17**	.32*	.41	–.10	2.03
Eye contact	.05	.34**	–.08	–.06	.16	.52**	–.16	1.30
Tempo/pace	–.31***	.24*	–.66***	–.22***	.10	–.19	–.05	–.81

Note. $N = 181$. E = Extraversion; A = Agreeableness; C = Conscientiousness; N = Neuroticism; O = Openness; EC = ego-control; ER = ego-resiliency. Standardized regression coefficients, Gelman's *d*, are reported, calculated as the change in the normative accuracy slope for a two standard deviation change in the information indicators divided by the dyadic normative accuracy random effect estimate standard deviation, in order to make these estimates comparable to the effect sizes for conditions (Gelman, 2008).
***$p < .001$. **$p < .01$. *$p < .05$. $^{+} < .10$.

The overall ANOVAs were statistically significant for each type of information: thoughts and feelings: $F(3, 112) = 40.06$, $p < .001$; specific behaviors: $F(3, 112) = 33.46$, $p < .001$; and general behaviors: $F(3, 112) = 79.57$, $p < .001$ (see Figure 1).

Tukey's honestly significant difference (HSD) post hoc tests were used to compare pairs of conditions. More information about thoughts and feelings was available in the Thoughts and Feelings condition ($M = 3.44$, $SD = .45$) than in all other conditions (Behavior Discussion 1: $M = 2.99$, $SD = .47$; Behavior Discussion 2: $M = 3.01$, $SD = .52$; Behaviors: $M = 2.04$, $SD = .53$, all $ps < .01$). More information about specific behaviors was available in the Behavior Discussion

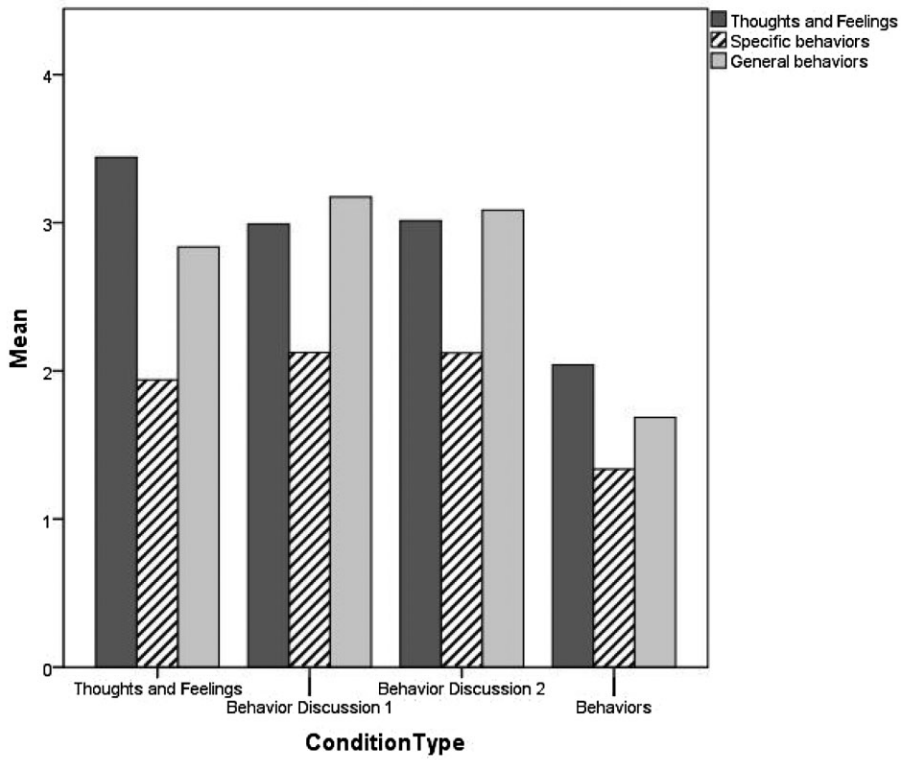

Figure 1 Amount of information available as a function of condition type.

conditions (Behavior Discussion 1: $M = 2.12$, $SD = .38$; Behavior Discussion 2: $M = 2.12$, $SD = .40$) than the Behaviors condition ($M = 1.34$, $SD = .24$), $ps < .001$, but the amount of information about specific behaviors did not differ in the Behavior Discussions and Thoughts and Feelings conditions ($M = 1.94$, $SD = .32$, $ps > .17$). More information about general behaviors was also available in the Behavior Discussion conditions (Behavior Discussion 1: $M = 3.17$, $SD = .43$; Behavior Discussion 2: $M = 3.08$, $SD = .40$) than the Behaviors condition ($M = 1.69$, $SD = .36$), $ps < .001$, and more information about general behaviors was also available in the Behavior Discussion 1 condition than the Thoughts and Feelings condition ($M = 2.84$, $SD = .44$), $p = .01$. However, the amount of information about general behaviors did not differ in the Behavior Discussion 2 and Thoughts and Feelings conditions, $p = .11$. Therefore, the manipulation of the type of information that was available in each condition was only partially successful. The Thoughts and Feelings condition was the most successful, in that more information about thoughts and feelings was available in this condition than in any other condition. The Behaviors condition was interesting because less of all types of information was available in this condition than in all other conditions, although the pattern of information availability was consistent with the Thoughts and Feelings condition. Even though the manipulation was only partially successful, it is still informative to compare accuracy across the experimental conditions to learn whether being instructed to discuss different kinds of information influenced the accuracy of judgments following the interaction.

Information Condition as a Moderator of Accuracy

On average and across conditions, for all of the personality traits, judges achieved significant levels of distinctive accuracy ($bs = .12$ to $.38$, all $ps < .001$) and normative accuracy ($bs = .49$ to 1.12, all $ps < .001$; see Table 2, first two columns of results).

Distinctive Accuracy. Distinctive accuracy was significantly higher in the Thoughts and Feelings condition than in the Behaviors condition for judgments of the full Big Five, Extraversion, and ego-control, all $ps < .01$ (see Table 2). Thus, dyads instructed to talk about thoughts and feelings achieved greater distinctive accuracy than dyads who engaged in behaviors for some traits, but distinctive accuracy did not differ for dyads instructed to talk about thoughts and feelings versus dyads instructed to talk about behaviors. Because participants in the Thoughts and Feelings condition, as compared to participants in the Behavior Discussion conditions, revealed more information about thoughts and feelings but about the same amount of

information about behaviors, it is possible that the information about behaviors compensated for having less information about thoughts and feelings.

Next, we examined whether there was an advantage to being instructed to talk about one's behaviors versus actually engaging in behaviors by changing the reference group in the analysis to the Behaviors condition. Greater distinctive accuracy was indeed achieved in both Behavior Discussion conditions compared with the Behaviors condition for judgments of the full Big Five, Extraversion, and ego-control, all ps < .05. Of note, distinctive accuracy did not reach significant levels in the Behaviors condition for judgments of Extraversion, Agreeableness, Conscientiousness, ego-control, or ego-resiliency, all ps > .12. Based on this pattern of results, there is evidence that learning about thoughts and feelings or behaviors based on a discussion is more useful for distinctive accuracy for some traits than is directly observing behaviors.

Normative Accuracy. Normative accuracy was only significantly higher in the Thoughts and Feelings condition for one comparison: with the Behavior Discussion 2 condition on judgments of ego-control, $p < .01$ (see Table 2). Thus, there is little evidence that impressions are more normatively accurate when people are instructed to discuss their thoughts and feelings than when they are instructed to discuss or engage in behaviors. Normative accuracy did not differ for any traits when comparing the Behavior Discussion conditions to the Behaviors condition.

Available Information as a Moderator of Accuracy

We next examined the extent to which the availability of the different types of information, as coded by observers, was related to distinctive and normative accuracy. See Table 1 for descriptive statistics for all codings of available information.

Distinctive Accuracy. The amount of information about thoughts and feelings was related to significantly higher levels of distinctive accuracy for Openness and ego-control, ps < .01 (see Table 3). The amount of information about both specific and general behaviors was related to significantly higher levels of distinctive accuracy on all Big Five traits, Extraversion, Neuroticism (only for specific behaviors), Openness, and ego-control (only for general behaviors), all ps < .05. Thus, there is evidence that talking about thoughts and feelings as well as behaviors is associated with greater distinctive accuracy, but talking about behaviors appears to be even more strongly and consistently associated with greater distinctive accuracy.

Two other types of information were also consistently associated with greater distinctive accuracy: discussing current life events and one's own personality (see Table 3). There was also surprising evidence that being more expressive was associated with lower distinctive accuracy for all Big Five traits, Conscientiousness, and Neuroticism.

Normative Accuracy. More information about thoughts and feelings was associated with significantly greater normative accuracy for all Big Five traits, Extraversion, Conscientiousness, Neuroticism, Openness, and ego-resiliency. Information about specific and general behaviors was associated with lower normative accuracy for all Big Five traits, Agreeableness, and Conscientiousness, and with higher normative accuracy for Extraversion, all ps < .05 (see Table 4).[6] A similar pattern of results was found for the other information indicators, such that more information was associated with higher normative accuracy for Extraversion but lower normative accuracy for the full Big Five and other specific traits (most consistently, Agreeableness and Conscientiousness).

Providing more information about relationships, being more nonverbally expressive, providing more personal information, and having a faster tempo/pace were the most consistent predictors of lower normative accuracy, whereas talking about past events and more assigned and additional topics were the most consistent predictors of higher normative accuracy.

DISCUSSION

Evaluation of Hypotheses

Hypothesis 1 was that verbal information about thoughts, feelings, and behaviors would contribute to higher distinctive accuracy and lower normative accuracy; this hypothesis received the most empirical support. Distinctive accuracy was higher for the Thoughts and Feelings and Behavior Discussion conditions compared to the Behaviors condition for some traits. Distinctive accuracy did not differ between the Thoughts and Feelings condition and the Behavior Discussion conditions, supporting the idea that information about thoughts and feelings is at least as useful as information about behavior. Distinctive accuracy was also higher for many of the traits when more information about specific behaviors and general behaviors was available, based on the objective codings.

When looking at normative accuracy, the predictions were not supported across the experimental conditions, but were partially supported when looking at coded information as a moderator. When significant effects emerged, normative accuracy tended to be lower when more information about thoughts and feelings, specific behaviors, and general behaviors was available. Thus, there is some evidence that normative accuracy tends to decrease as higher-quality information becomes available, but there were exceptions to this, such as in the case of Extraversion, as will be discussed in more detail below.

Hypothesis 2 was that information about thoughts and feelings would be the most strongly related to increases in distinctive accuracy and decreases in normative accuracy. Support for this hypothesis was found for distinctive accuracy when the Thoughts and Feelings condition was compared to the Behaviors condition, but not when compared to the Behavior Discussion conditions. Furthermore, verbal information about behaviors was more consistently and slightly more strongly

related to higher distinctive accuracy and lower normative accuracy than information about thoughts and feelings, which is contrary to this prediction. Interestingly, the benefits of verbal information about behaviors for distinctive accuracy appeared to apply to both high and low observability traits, despite the fact that thoughts and feelings information should theoretically be more relevant to more internal traits than external, observable traits.

Hypothesis 3 was that the Behaviors condition would be the most strongly related to increases in distinctive accuracy and decreases in normative accuracy. There was no evidence to support this hypothesis, even for highly observable traits for which behavioral information might be the most relevant. Specifically, when compared to the Behaviors condition, both Behavior Discussion conditions resulted in higher levels of distinctive accuracy for the Big Five traits combined, Extraversion, and ego-control. This is particularly interesting given that in one of the Behavior Discussion conditions, participants discussed their behaviors in activities very similar to the activities that participants in the Behaviors conditions engaged in. Thus, it is not just that those types of behaviors are not diagnostic of personality, but instead it seems that talking about them is more useful for accurately judging personality than engaging in them. One reason this might be the case is that any given behavior may not be perfectly predicted by one's personality, whereas when one discusses one's behaviors one is able to mentally aggregate across a number of different experiences, thereby providing more reliable information about typical behavior (e.g., Borkenau et al., 2004; Epstein, 1983).

Hypothesis 1 was based on theory, Hypothesis 2 was based on a previous empirical finding, and Hypothesis 3 was based on philosophy. It is interesting to note that the theoretical hypothesis was the most strongly supported.

Implications

This pattern of findings has several interesting implications. First, the kind of accuracy that is being examined is important. The majority of the statistically significant effects that were found when comparing experimental conditions were for distinctive accuracy rather than normative accuracy. This suggests that people are learning more individuating information in all three discussion conditions than in the Behaviors condition, and they are using this individuating information to make more accurate judgments about how their partners are unique. In fact, on average, participants in the Behaviors condition did not achieve significant levels of distinctive accuracy on most of the individual traits. However, the discussion conditions do not give judges a consistent advantage for normative accuracy, and participants in all conditions achieved significant levels of normative accuracy on all traits.

A second implication of the findings from the experimental manipulation is that being instructed to talk about thoughts and feelings does not lead to consistently different levels of either distinctive or normative accuracy than does being instructed to talk about behaviors. This finding, which implies that information quality is about equal for thoughts and feelings as for behaviors, is consistent with Hypothesis 1 but inconsistent with Hypothesis 2. However, based on the codings of available information, it appears that information about behaviors is higher quality than information about thoughts and feelings.

Other types of information also appear to be useful for increased distinctive accuracy, including information about current events and the targets' own personalities. Additionally, more expressive targets were judged with lower levels of distinctive accuracy. It is unclear why expressiveness would decrease accuracy, and this finding should be replicated in future work.

The most interesting aspect of the results for normative accuracy is that information quality is related to both higher and lower levels of this type of accuracy. The majority of the statistically significant associations were negative, such that greater information quality was associated with lower normative accuracy. Given the positivity of the normative profile (Edwards, 1957), one implication of these results is that obtaining higher-quality information may actually result in more negative personality impressions. This is similar to findings that more motivated perceivers form more distinctively accurate impressions, but this comes at the cost of less normative and therefore less positive impressions (Biesanz & Human, 2010).

However, despite this general negative trend, there were also many positive associations between information quality and normative accuracy. The positive effects were primarily for Extraversion, and the negative effects were primarily for Agreeableness and Conscientiousness. It is not clear why higher-quality information would increase normative accuracy for one trait and decrease it for another, and this should also be examined further. One possibility is an interaction with the observability of the trait being judged. Extraversion is the most observable trait of the Big Five (Funder & Dobroth, 1987), and this was the trait that had the most positive relations between amount of information available and normative accuracy. Thus, obtaining higher-quality information about highly observable traits may actually enhance utilization of normative information or result in more positive impressions, whereas higher-quality information about less observable traits seems to decrease the utilization of such normative, positive information.

Notably, information about thoughts and feelings and specific behaviors was associated with increases in both distinctive and normative accuracy for Extraversion and Openness. Although it seems like distinctive and normative accuracy should always be inversely related, they are in fact statistically independent of one another (Biesanz & Human, 2010; Human & Biesanz, 2011b), and in this instance, both types of accuracy for Extraversion increased when more information was available.

Overall, these findings have interesting implications for the types of questions that should be asked and the types of situ-

ations that should be constructed when attempting to get to know someone, such as during job interviews or on first dates. For example, these findings provide support for the common practice of "behavioral interviews" (Motowidlo et al., 1992), in which interviewees are asked to describe their behaviors in previous situations, by demonstrating that talking about one's behaviors does indeed promote more distinctively accurate impressions. Furthermore, information about behaviors was associated with higher distinctive accuracy for Agreeableness and Conscientiousness, which are consistently associated with occupational performance, success, and attainment (Ozer & Benet-Martínez, 2006). On the other hand, impressions regarding some traits may become less normative and less positive when such information is available, which is important to keep in mind when evaluating others, such as job candidates. Additionally, asking someone to discuss his or her typical thoughts, feelings, and behaviors across situations may lead to more distinctively accurate personality judgments because this kind of information is what should be best predicted by traits, and therefore this information may be easier to utilize when making trait judgments. In contrast, in the Behaviors condition, the behaviors were all single instances and could have varied from typical behavior. The judges would not know that the behavior was atypical unless the person said so. For this reason, cues in the Behaviors situation may have lower relevance and be more difficult to utilize. It will be important to examine whether the same pattern of effects emerges in different social contexts, including job interview situations and more naturalistic settings.

CONCLUSION

It appears that Plato got it backwards, and an hour of conversation is more useful for learning about a person than an hour of play. If you want to accurately determine how people are unique, you should encourage them to tell you about both their specific behaviors and how they generally behave in certain situations.

Notes

1. Three additional groups were excluded from the analyses due to failure to follow the research protocol.
2. Satisfaction with life (Diener, Emmons, Larsen, & Griffin, 1985) was also assessed, but it was not used in the current analyses.
3. The numbers of assigned and additional topics discussed were counted by the coders and were not assessed on a 4-point scale.
4. When only the acquaintance ratings were used as the accuracy criteria, the results were similar to the reported findings, and the overall conclusions did not change. This similarity demonstrates that the effects are based on more than shared method variance that could result from the targets providing their self-ratings after the interaction in which they revealed certain information about themselves.
5. This effect is referred to in this way to clearly distinguish it from other *d*-family effect sizes.
6. The larger *d* values for ER appear to be due to there being fewer individual differences in normative accuracy for ER. It is unclear why there would be less variability in normative accuracy for ER relative to the other traits.

References

Andersen, S. M. (1984). Self-knowledge and social inference: II. The diagnosticity of cognitive/affective and behavioral data. *Journal of Personality and Social Psychology*, **46**, 294–307.

Beer, A., & Brooks, C. (2011). Information quality in personality judgment: The value of personal disclosure. *Journal of Research in Personality*, **45**, 175–185.

Biesanz, J. C. (2010). The social accuracy model of interpersonal perception: Assessing individual differences in perceptive and expressive accuracy. *Multivariate Behavioral Research*, **45**, 853–885.

Biesanz, J. C., & Human, L. J. (2010). The cost of forming more accurate impressions: Accuracy-motivated perceivers see the personality of others more distinctively but less normatively than perceivers without an explicit goal. *Psychological Science*, **21**, 589–594.

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

Block, J., & Kremen, A. M. (1996). IQ and ego-resiliency: Conceptual and empirical connections and separateness. *Journal of Personality and Social Psychology*, **70**, 349–361.

Borkenau, P., & Liebler, A. (1995). Observable attributes as manifestations and cues of personality and intelligence. *Journal of Personality*, **63**, 1–25.

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, **86**, 599–614.

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

Colvin, R. C., & Funder, D. C. (1991). Predicting personality and behavior: A boundary on the acquaintanceship effect. *Journal of Personality and Social Psychology*, **60**, 884–894.

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

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, **49**, 71–75.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden.

Epstein, S. (1983). Aggregation and beyond: Some basic issues on the prediction of behavior. *Journal of Personality*, **51**, 360–392.

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

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

Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology*, **52**, 409–418.

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.

Furr, R. M. (2008). A framework for profile similarity: Integrating similarity, normativeness, and distinctiveness. *Journal of Personality*, **76**, 1267–1316.

Gangestad, S. W., Simpson, J. A., DiGeronimo, K., & Biek, M. (1992). Differential accuracy in person perception across traits: Examination of a functional hypothesis. *Journal of Personality and Social Psychology*, **62**, 688–698.

Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, **27**, 2865–2873.

Gifford, R. (1994). A lens-mapping framework for understanding the encoding and decoding of interpersonal dispositions in nonverbal behavior. *Journal of Personality and Social Psychology*, **66**, 398–412.

Human, L. J., & Biesanz, J. C. (2011a). Target adjustment and self-other agreement: Utilizing trait observability to disentangle judgeability and self-knowledge. *Journal of Personality and Social Psychology*, **101**, 202–216.

Human, L. J., & Biesanz, J. C. (2011b). Through the looking glass clearly: Accuracy and assumed similarity in well-adjusted individuals' first impressions. *Journal of Personality and Social Psychology*, **100**, 349–364.

International Personality Item Pool: A scientific collaboratory for the development of advanced measures of personality traits and other individual differences. (n.d.). Retrieved from http://ipip.ori.org/

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

Letzring, T. D., Block, J., & Funder, D. C. (2005). Ego-control and ego-resiliency: Generalization of self-report scales based on personality descriptions from self, acquaintances, and clinicians. *Journal of Research in Personality*, **39**, 395–422.

Letzring, T. D., Wells, S. M., & Funder, D. C. (2006). Quantity and quality of available information affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology*, **91**, 111–123.

Motowidlo, S. J., Carter, G. W., Dunnette, M. D., Tippins, N., Werner, S., Burnett, J. R., et al. (1992). Studies of the structured behavioral interview. *Journal of Applied Psychology*, **77**, 571–587.

Murray, H. A. (1943). *Thematic Apperception Test manual*. Cambridge, MA: Harvard University Press.

Naumann, L. P., Vazire, S., Rentfrow, P. J., & Gosling, S. D. (2009). Judgments of personality based on physical appearance. *Personality and Social Psychology Bulletin*, **35**, 1661–1671.

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, **57**, 401–421.

Rogers, K. H., & Biesanz, J. C. (2013). Knowing vs. liking: Separating normative knowledge from social desirability in personality impressions. Manuscript submitted for publication.