



Judging personal values and personality traits: Accuracy and its relation to visibility



Jennifer S. McDonald*, Tera D. Letzring

Idaho State University, Department of Psychology, 921 S. 8th Street, Stop 8112, Pocatello, ID 83209, United States

ARTICLE INFO

Article history:

Received 1 April 2016

Revised 26 September 2016

Accepted 17 October 2016

Available online 19 October 2016

Keywords:

Judgmental accuracy

Personality judgment

Personal values

Visibility

ABSTRACT

Personality judgment research typically focuses on traits, and visibility of traits often predicts judgmental accuracy. However, visibility and accuracy of values are not well-explored. It was predicted that visibility and accuracy would be positively related, and values would be judged more accurately than traits. In groups of 3 acquaintances, 204 undergraduates completed self-report measures of traits and values, and other-reports of acquaintances. A separate sample of 247 participants assessed visibility. Values were rated as more visible than traits, but traits were judged more accurately than values. Correlations between visibility and item-level accuracy for both values and traits were small and non-significant. The ease with which people think values and traits can be judged may differ from the accuracy of judgments.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

You are always communicating, even when you are silent. This statement refers to how we share who we are with others, whether we are speaking or not. Individuals continually convey messages about who they are and who they think they are, through their personal appearance, behaviors, words, emotional expressions, and many other cues. The ability to accurately judge messages sent by others through these cues is important, as it contributes to an awareness of who best to include or avoid in many domains of life.

The historical background on person perception suggests that in years past, individuals were judged by others based on their principles, such as loyalty or honesty, but in current times principles have given way to a reliance on personality (Nicholson, 2003). However, some researchers are advocating for an empirical investigation of character from a Personality Psychology perspective (Fleeson, Furr, Jayawickreme, Meindl, & Helzer, 2014), toward which the current study may contribute further information.

An individual's personality is comprised of many different features, which include personal values, personality traits, attitudes, and beliefs, just to name a few. Certain features of a person, such as values, may describe self-beliefs of an individual; while other features, such as traits, may speak to who others perceive them to be (McAdams, 1995). A pertinent question involves all of these features of personality: Is it possible to truly know, objectively,

who a person is? Numerous studies have attempted to answer this question using inter-peer agreement or consensus (e.g., Beer & Watson, 2008; Costa & McCrae, 1988; Funder, 1995; McAdams, 1995), which refers to how well two people agree on what a specific person is like. Of all the features representing a person, personality traits have frequently been the focus when examining accuracy of personality judgments. In particular, the broad trait domains of the Five Factor Model (extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience) are most commonly employed. The current study examines how personal values, as a feature of personality and as a feature of character, compare to traits in terms of visibility the accuracy with which they are judged by others.

One of the most widely accepted definitions of traits are the unique ways individuals tend to exhibit enduring patterns of thoughts, feelings, and actions (McCrae & Costa, 1990). The five factor model (FFM, Goldberg, 1993; McCrae & John, 1992) is a way to present the broad fundamental aspects under which specific traits may be organized (Vecchione, Alessandri, Barbaranelli, & Caprar, 2011).

Although personality traits are useful in describing and predicting behavior, they cannot explain all aspects of personality. Pervin (1994) expressed that traits are unable to completely account for all behavior, because while they have the capacity to describe patterns and consistencies in behavior, they cannot provide an explanation of the origin of those patterns and consistencies. Incorporating values into the study of personality and accuracy

* Corresponding author.

E-mail address: mcdojenn@isu.edu (J.S. McDonald).

may be useful in filling in some of the gaps left by traits, by helping to explain the “why” of behavior.

Personal values represent motivations that drive people to make decisions in the manner in which they do (Bilsky & Schwartz, 1994) and are “goals that act as guiding principles in one’s life” (Schwartz, 1992, p. 46). The ten value domains from the Schwartz Value Survey have been found to be fairly stable across situations, time, and culture in many studies (Roccas, Sagiv, Schwartz, & Knafo, 2002; Schwartz, 1992; Schwartz & Sagie, 2000; Schwartz, Melech, Lehmann, Burgess, Harris, & Owens, 2001; Vecchione, Alessandri, Barbaranelli, & Caprar, 2011). It is the stability across situations that allows values to fall into the category of personality. Values express stability by representing goals across situations, however, they also fluctuate in importance by person and situation (Bilsky & Schwartz, 1994; Schwartz, 1992; Schwartz & Bilsky, 1987, 1990).

While it is recognized that traits and values are both helpful in explaining personality, they are different constructs. The same phrases or words can be construed as a trait, or as a value, yet have very different meanings, such as obedience or success. One fundamental difference between values and traits is that traits are seen as descriptions of observed patterns of behavior, whereas values are used by individuals to judge the desirability of behavior, people, and events (Bilsky & Schwartz, 1994). Although they are different constructs, it may be possible to accurately judge the values of another person in a similar way to how one judges traits.

Personality judgment is often examined using the Realistic Accuracy Model (RAM; Funder, 1995). The model acknowledges that many sources of information are needed in order to achieve the highest level of accuracy. According to RAM, accuracy is possible following completion of the stages of relevance, availability, detection, and utilization of behavioral cues involving real people in realistic situations. Relevance refers to the cues that are appropriate to the trait being judged. Some situations or contexts are more conducive to the expression of certain trait-relevant behavior than others. The availability stage means that the cues must be present in a way that makes it possible for the judge to perceive them. The importance of the relevance and availability stages to accuracy has been emphasized by Letzring (2008), who found that the warm and agreeable personality of a judge can create comfortable situations that bring out both relevant and available cues from the target. Once made available, the judge must detect the behaviors relating to the trait being judged. Some available behaviors may be more detectable than others, such as the level of talkativeness versus subtle facial expressions of the target. Finally, the judge must have the capacity to utilize the cues correctly to make an accurate judgment that the person possesses high or low levels of a trait.

The model also proposes that accurate personality judgments are best achieved with the presence of the following four moderators: a good judge, a good target, good traits, and good information. The good judge refers to the idea that some people are more accurate judges than others. For example, judges with a stronger orientation toward interpersonal relationships are more accurate than those who do not view interpersonal relationships as important (Vogt & Colvin, 2003). A good target means that some people are easier to judge than others, such that they provide more relevant and available cues to the judge. Overall, good targets share similar characteristics that pertain to psychological adjustment, higher social status, and socialization of skills which promote judgability (Colvin, 1993; Human & Biesanz, 2013). The moderator of good traits suggests that some traits are easier to judge than others. For example, extraversion is considered a good trait as it is more accurately judged due to the high number of relevant and available cues associated with it. Lastly, good information consists of two aspects: quantity and quality. Information quantity

refers to the idea that knowing someone longer and seeing him or her more often will provide more information about that person. Information quality reflects that certain relationship contexts or the type of subjects discussed provide better information than others. For example, discussing behaviors provided higher quality information and contributed to higher distinctive accuracy compared to discussing thoughts and feelings, or engaging in actual behaviors (Letzring & Human, 2013).

Accuracy of personality judgment is also divided into two components: normative and distinctive accuracy. Normative accuracy is the degree to which a judge’s ratings of a target correspond with the ratings of the average person, while distinctive accuracy is the ability of the judge to accurately rate how a person is unique or different from the average person.

Some characteristics of traits are related to accuracy of personality judgment, including observability or visibility and evaluativeness¹ (the degree to which a trait is desirable or favorable) (Funder & Drobth, 1987; John & Robins, 1993). For instance, the visibility of a trait, which represents the presence of relevant and available cues of that trait, is predictive of the accuracy with which the trait can be judged (Funder & Drobth, 1987). This characteristic of good traits can also be applied to personal values, because some values may also be more visible than others. Although they may be more favorable than traits, values may also be similar to traits in that they vary somewhat in evaluativeness and the accuracy with which they are judged may be impacted by who is doing the judging: self or other.

As mentioned previously, some studies have found that the more visible a trait is, the more accurately it is judged (Funder & Drobth, 1987; John & Robins, 1993; Paunonen, 1989). However, for some traits, the findings linking visibility and accuracy have been mixed. Funder and Drobth (1987) found that items indicative of extraversion were significantly more visible than those associated with other traits, while items related to neuroticism were rated as the least visible of all the traits. As expected, the trait items with the highest level of visibility were also judged more accurately. However, contrary to expectations given the overall relationship between visibility and accuracy, only the neuroticism items that were rated less visible were also judged the least accurately. Several of the other trait items with the lowest level of visibility were not always judged the least accurately.

In addition, John and Robins (1993) found extraversion to be judged with the highest accuracy, which could be explained by its high visibility. However, in the same study, emotional stability (neuroticism) was not found to have low agreement or low visibility, which does not support previous findings by Funder and Drobth (1987). It appears from these inconsistent findings that high visibility is related to accuracy of traits, but that low visibility may not always be related to accuracy.

The mixed findings of visibility and accuracy may be due to several factors, such as the type of questions used to assess visibility, how accuracy is measured, and the acquaintance level of the targets and judges. Some studies not finding a relationship between visibility and accuracy have assessed visibility with one unipolar question (John & Robins, 1993; Paunonen, 1989). Funder and Drobth’s (1987) study incorporated the visibility of several bi-polar dimensions which may account for the link found between visibility and accuracy. The current study was conducted in a similar manner to Funder and Drobth’s and sought to replicate their findings between visibility and accuracy for traits, and in addition, to examine the same relationship between visibility

¹ Observability and visibility have both been used to refer to the same characteristic of traits. Visibility will be used in this manuscript to refer to this characteristic. Likewise desirability and favorability have both been used to refer to the same characteristic of traits. Favorability will be used in this manuscript to refer to this characteristic.

and judgmental accuracy for personal values.

Visibility of values ratings may be especially high for people who have known the individual being judged for some time, due to the acquaintanceship effect (the longer people have known each other, the more information they have obtained about one another and the more accurately they are able to judge each other; Colvin & Funder, 1991). A longer acquaintanceship is also linked to increased interjudge agreement of personality traits (Colvin & Funder, 1991; Paunonen, 1989). However in one study the relationship between accuracy of personality judgment and visibility was only found for those less acquainted, and not for those in close relationships (Paunonen, 1989). Highly acquainted people are very familiar with one another across many different contexts so they may accurately judge the traits of one another regardless if that trait is visible or not.

The favorable nature of values may also contribute to an increase in visibility. Values are typically conceptualized as being positive, thus people may actively endeavor to inform others of what their values are and this may improve the relevance and availability stages of RAM and in turn lead to greater accuracy (Roccas et al., 2002). Although, the favorability of traits has been found to be related to the expected frequency of traits (Funder & Dobroth, 1987; Rothbart & Park, 1986), highly evaluative traits are usually not judged accurately due to the fact that there are some traits that people attempt to hide from others, and in some cases, from themselves (John & Robins, 1993; Vazire, 2010). Thus, values may yield a stronger relationship between visibility and accuracy than traits.

To our knowledge, accuracy of judgments of values has only been examined in one study (Dobewall, Aavik, Konstabel, Schwartz, & Realo, 2014) and the relationship between visibility and accuracy of judgments of values has not been studied specifically. However, research has examined the link between an individual's values and the behaviors associated with those values (Bardi & Schwartz, 2003; Pozzebon & Ashton, 2009). It was found that self-reports of the stimulation and tradition value types were the best predictors of self-reported and peer-reported behaviors associated with those value types. Values and traits have also been shown to be almost equally strong in predicting value-relevant behaviors, with personality traits being slightly stronger predictors of behavior overall (Pozzebon & Ashton, 2009). However, behaviors relevant to some values (e.g., tradition) were better predicted by values, and behaviors relevant to other values (e.g., self-direction, stimulation, hedonism, and power) were better predicted by traits. Results from these studies indicate that there is a relationship between some of the personal values and observable cues related to those values.

It has been established that there are many similarities and differences between traits and values. In order to parse out their unique contribution to the study of accuracy of personality judgment, the current study sought to determine if the level of visibility of values yields the same relationship with accuracy as that of the level of visibility of traits. If values can be judged accurately, then they may be utilized as a way to begin to accurately judge another's motivations. It would also be helpful to know if accuracy of judging values is related to visibility of those values. If an individual's motivation can be known, it is possible that the person's behavior may be better understood and perhaps more accurately predicted. The information derived from accurately judging another's motivation could potentially be beneficial in selecting and motivating employees and college students, as well as selecting friends and romantic partners. To begin this process of understanding an individual's motivation, it is helpful to examine the visibility of values as a first step.

The current study sought to replicate and extend the findings of Funder and Dobroth (1987) to determine whether values would yield greater visibility than traits and if the visibility of values

had a similar relationship with accuracy as that of traits. Self-other agreement was also examined for the higher-order values as well as value types to compare to the findings of Dobewall et al. (2014). Values were examined as good constructs by measuring the visibility and interpersonal accuracy of the value items. It may be that values are more visible than traits due to their greater favorability, which may increase the availability and detection stages of RAM, leading to increases in judgmental accuracy.

The following hypotheses were tested.

Hypothesis 1. Overall, personal values would be rated as more visible than personality traits due to the greater availability of value-relevant cues.

Hypothesis 2. Personal values would be judged more accurately than personality traits, while holding visibility constant. This hypothesis is based on the motivational nature and favorability of values, both of which are assumed to be higher than that of traits.

Hypothesis 3. Similarly to traits, more visible values would be judged more accurately than less visible values. More specifically, based on the findings of Bardi and Schwartz (2003), stimulation and tradition value types were expected to be the most visible and yield the greatest accuracy of all the personal values.

2. Method

2.1. Participants

Two-hundred four undergraduate students were recruited for this study (62% female, 75% Caucasian, 14% Latino, 4% African-American, 4% Asian/Pacific Islander, 1% Native-American, and 2% Middle Eastern). Age of participants had a mean of 20.07 ($SD = 2.22$), within a range of 18–30 years. Religious affiliation was assessed by a free-response question and put into categories by the researcher. Twenty-six percent of participants were LDS, 17% Catholic, 23% Christian, 4% Atheist, 23% no religion, and 7% other. Sixty-eight main participants were recruited through the Psychology department's research participant pool and 136 additional participants (acquaintances of the main participants) were invited by the main participants. The median relationship length for all acquaintances was 2 years. Participants identified the types of relationships they had with each of the targets and were able to choose more than one category. Eighty-six percent of the total participants rated themselves as friends, 14% roommates, 28% classmates, 8% co-workers, 4% siblings, 8% significant others, and 4% as other.²

Main participants were reimbursed with one unit of course credit per half hour of participation (the study took about one hour). Acquaintances were given the choice between receiving course credit if they were enrolled in a psychology course where the credit was valid, or a buy-one-get-one-free coupon for a local coffee establishment.

Two-hundred forty-seven additional participants who were not included in the accuracy study were recruited through the Psychology Department participant pool to assess the visibility of traits and values ($M_{age} = 21.73$, $SD_{age} = 5.80$, 78% female, 83% Caucasian, 8.1% Latino, 1.6% African-American, 3.2% Asian/Pacific Islander, 1.6% Native American, and 2.5% other).

² The total percentage is greater than 100% because participants were allowed to select more than one category.

2.1.1. Statistical power

Sample sizes were determined using G-Power statistical software. For the first hypothesis, which is that visibility of personal values would be greater than visibility of personality traits, there is a paucity of research comparing the subjective visibility of traits to other characteristics from which to determine expected effect sizes. Although differences in visibility between values and traits have not been specifically examined, it was expected that 128 total participants would have achieved sufficient power for an 80% chance of detecting a moderate effect size of $d = 0.5$ at $p = 0.05$ based on an independent-samples t -test comparing visibility between values and traits. Previous research examining observability between traits found that the Big Five traits differed in terms of observability with an effect size of $d = 0.40$ (John & Robins, 1993). If the aim had been for an effect size of 0.40 a total of 200 participants would have been required, which was surpassed in the current visibility sample.

Hypothesis 2, which is that accuracy of values would be greater than accuracy of traits, was tested in two ways. First, using item-level correlations, power is based on a certain number of items. In this study there was a total of 156 trait and value items which was expected to yield sufficient power for a 98% chance of detecting a small effect size of $d = 0.3$ at $p = 0.05$ based on an independent samples t -test between the means of overall accuracy.

Hypothesis 2 was also tested using profile correlations. It was initially expected that 100 total participants would have achieved sufficient power for an 80% chance of detecting a moderate effect size of $d = 0.5$ at $p = 0.05$ based on a paired-samples t -test comparing accuracy that each participant obtained for judgments of traits and values. Research examining the difference between accuracy of traits and accuracy of values published after the original power analysis was conducted, found an effect size of $d = 0.29$ (Dobewall et al., 2014). Therefore, in the current study, power would be 0.99 to detect a similar effect size of $d = 0.29$ based on a paired-samples t -test of 204 participants.

Hypothesis 3 examined the relationship between accuracy and visibility of personal values. The sample size for the correlation between accuracy and visibility is based on the number of items in the measures, so again it is not possible to increase power with a large sample size. However, Funder and Dobroth (1987) found an overall correlation between self-other agreement and visibility of traits of $r = 0.39$. John and Robins (1993) found similar overall correlations between observability and peer-peer agreement for traits, r 's = 0.36 and 0.37 in two different studies. The correlations between observability and self-peer agreement for traits were $r = 0.38$ and $r = 0.50$ for Study 1 and 2 respectively, so we can expect overall correlations between visibility and accuracy to be around the 0.40 level. In addition, the visibility and accuracy correlation for values cannot be compared with the visibility and accuracy correlation for traits by tests of independent or dependent correlations because the ratings are not independent nor are they dependent in the typical way. To overcome this issue, the focus of the analyses was on effect sizes and confidence intervals.

2.2. Measures

2.2.1. Personality traits

Personality traits were assessed using the International Personality Item Pool 100-item version of the NEO-PI-R domains (IPIP NEO-PI domains). This measure consisted of 20 items per trait domain with responses based on a 5-point scale ranging from 1 (*very inaccurate*) to 5 (*very accurate*). Questions are both positively and negatively scored. The IPIP version of the NEO-PI-R has good established reliability with Cronbach's alpha ranging from 0.85 (for agreeableness) to 0.91 (for extraversion and neuroticism; International Personality Item Pool, n.d.). Using Cronbach's alpha, the self and other personal-

ity ratings in the current study yielded good reliabilities for extraversion (0.93), agreeableness (0.94), conscientiousness (0.95), neuroticism (0.92), and openness to experience (0.91).

2.2.2. Personal values

Personal values were measured using the Schwartz Values Survey (SVS; Schwartz, 1992). The measure consists of two lists of value items. The first list has 30 items that describe possible end-states, such as *Meaning in life (a purpose in life)*. The second list contains 26 items that describe possible ways of behaving to achieve the end states from the first list, such as *Healthy (not being sick physically or mentally)*. Each value is rated on importance "as a guiding principle in my life" on an asymmetrical 9-point scale from -1 (*Opposed to my values*), 0 (*Not important*), 3 (*important*), 6 (*very important*), and 7 (*of supreme importance*). The asymmetrical scale is used to indicate which values are perceived as favorable or favorable. Each value type is measured with between three (hedonism) and eight (universalism) items. The SVS has adequate reliability with alphas for hedonism (0.64), power (0.72), achievement (0.72), stimulation (0.70), self-direction (0.60), universalism (0.72), benevolence (0.67), tradition (0.63), conformity (0.61), and security (0.61) (Roccas et al., 2002). Cronbach's alphas in the current study were acceptable for all value ratings: hedonism (0.80), power (0.86), achievement (0.86), stimulation (0.80), self-direction (0.83), universalism (0.89), benevolence (0.89), tradition (0.84), conformity (0.83), and security (0.83).

2.2.3. Visibility

Visibility was assessed by five of the eight dimensions developed by Rothbart and Park (1986); one dimension developed by Funder and Dobroth (1987): (*easiness to judge*), and an additional item created by the current researchers to assess how easy it would be for others to judge the trait/value in yourself (*ease to judge in yourself*); (see Appendix A). Alphas for the dimensions regarding traits are as follows: ease of imagining confirmed behaviors (0.70), ease of imagining disconfirmed behaviors (0.72), occasions for confirming behavior (0.81), occasions for disconfirming behavior (0.76), favorability (0.97) (Rothbart & Park, 1986), and easiness to judge in others (0.87) (Funder & Dobroth, 1987). Three of the original nine dimensions by Rothbart and Park (1986) (number of behavioral instances to confirm or disconfirm the trait and prevalence of the trait in the population) were not used due to low rater reliability found by Funder and Dobroth (1987). Although not measuring observability specifically, Funder and Dobroth (1987) referred to these dimensions of trait properties collectively as visibility. Other researchers have used these dimensions to rate visibility of the California Q-Set items (Funder & Dobroth, 1987), whereas the current study rated the visibility of traits from the IPIP NEO-PI-R 100-item measure, and values from the 56 items of the SVS. Eight sub-versions of the questionnaire were created to ensure participants were not fatigued by the number of item ratings. Five of the questionnaires consisted of 20 (out of 100) trait items, each rated on seven dimensions of visibility. Three of the questionnaires consisted of 18 or 19 value items (out of 56) rated on seven dimensions of visibility. Responses were made using a 9-point scale from 1 (*quite difficult*) to 9 (*quite easy*) for the ease of imagining confirmed or disconfirmed behavior, and the easiness to judge dimension. Other responses were adapted to fit the remaining items, such as 1 (*very unfavorably*) to 9 (*very favorably*) for the favorability dimension. In the current study alphas for the visibility dimensions were ease of imagining confirmed behaviors (0.90; 0.82), ease of imagining disconfirmed behaviors (0.92; 0.85), occasions for confirming behavior (0.89; 0.79), occasions for disconfirming behavior (0.88; 0.84), easiness to judge in others (0.89; 0.77), easiness to judge in self (0.88; 0.77), and favorability (0.84; 0.73) for values and traits respectively.

2.3. Procedure³

2.3.1. Acquaintance judgments

Participants arrived at the research laboratory together in groups of three, wherein everyone had known one another for at least six months. The use of acquainted groups is consistent with the procedure of Funder and Dobroth (1987) and provided the participants with real-world information on which to base judgments of personal values. Before beginning, a research assistant informed participants that the ratings they provided of their own traits and values, as well as those of others, would be kept confidential. Using data collection software on laptop computers located in different areas of the lab to ensure privacy, participants completed self-reports of traits assessed by the IPIP NEO-PI domains, and values assessed by the SVS. After making their self-ratings, participants rated their acquaintances' traits and values using other-report versions of the same IPIP NEO-PI domains and the SVS, and reported the length of acquaintanceship and degree of familiarity with each person. A question of "What is your relationship with this individual?" with fixed answers (e.g., family member, friend, romantic partner) was included to determine the context of the relationship. Responses for both acquaintances were made one at a time in a "round-robin" method. Participants were instructed to rate the person to their right and later to their left and entered the names of the targets into the software in order to continually reference who was being rated. Finally participants were asked to complete a brief demographics measure.

2.3.2. Visibility ratings

Consistent with the method used by Funder and Dobroth (1987), visibility ratings were made by a separate sample than those making accuracy ratings to avoid any influence on the self and acquaintance ratings. The separate sample also reduced the amount of fatigue experienced by participants.

Participants in the visibility portion of the study were recruited through the Psychology department's research pool and directed to an internet survey site to complete a questionnaire on either the visibility of values or traits in order to earn course credit. Participants had a choice to complete one of eight available questionnaires without knowledge of whether traits or values would be rated. For ratings of values, participants signed a digital consent form and read instructions on how to complete the study based on the instructions from the Rothbart and Park (1986) visibility study, yet adapted for values:

Personal values represent motivations that drive people to make decisions in the manner in which they do. For some values that people possess, judgments of whether someone has the value or not may be relatively easy. For other values, their presence or absence is much more difficult to judge. For example, consider the value "cares about the environment." Is it ordinarily rather easy to determine whether someone has this value or not? For each of the 19 personal value item terms on the following screens, please rate how difficult or easy you think it would be to judge another person.

The procedure for the ratings of traits was the same as for values, except the instructions for completing the measure were used directly from the Rothbart and Park (1986) study.

2.4. Analyses

There are several parts to the analyses. For each part, scores were calculated for both values and traits. To test Hypothesis 1,

visibility scores were calculated for values and traits so that overall visibility of the characteristics could be compared. For Hypothesis 2, accuracy scores were calculated in two different ways. First, accuracy scores were calculated at the item-level – meaning that raw scores were correlated across judge-target pairs for each item. These scores reflect the consistency between the ordering of self-ratings and acquaintance-ratings. The sample size for each correlation is the number of judge-target pairs. An advantage of this approach is that a stable accuracy score could be computed for each item due to the relatively large number of judge-target pairs. Second, accuracy scores were calculated at the profile level – meaning that raw scores were correlated across items for each judge-target pair. This approach allows for two accuracy scores to be calculated for each judge, one for traits and one for values, which then makes it possible to directly compare accuracy of values and traits. For Hypothesis 3, item-level accuracy scores for values and traits were again used because this made it possible to correlate accuracy scores with visibility scores for each item, and to then compare with these correlations with each other.

2.4.1. Addressing dependencies in the data

Due to participants being grouped together in triads and acting as both targets and judges, the problem of violating the assumption of independence of data arose. This issue was addressed through two different analytic strategies – first by analyzing the data using three separate subsets of the data, and second through the use of multilevel modeling.

Three subsets of the sample were created using one member of each triad as the target in each subset. For the first subset, the 68 main participants initially recruited for participation were treated as targets, and the two acquaintances recruited by these main participants were treated as judges. For the second subset, the acquaintances deemed Acquaintance As were treated as targets, and the main participants and Acquaintance Bs were treated as judges. In the third subset, only Acquaintance Bs were treated as targets, and the main participants and Acquaintance As were treated as judges. In all three subsets, the correlations between the two types of accuracy and visibility at the item-level were similar to those found when using the entire sample, and therefore results from the entire sample are presented in the Results section. Further, profile correlations of self-other agreement calculated within each subset showed the same pattern of results that was found using all 204 participants as both targets and judges. For this reason, results from the entire sample are again presented in the Results section. See Table S1 in the online supplemental materials for results from the analyses based on each subset.

The other analytic method used, multi-level modeling, does not specifically address the issue of dependency in the data, but it does account for the nested structure of the data. For correlations calculated at the profile-level, a multilevel model called the Social Accuracy Model (SAM, Biesanz, 2010) can be conducted to simultaneously estimate two components of accuracy: normative and distinctive accuracy. The analyses of fixed effects revealed that when the nested structure of the data was accounted for, judges achieved significant normative, $b = 0.95$, $p < 0.001$ and distinctive accuracy, $b = 0.18$, $p < 0.001$ of traits, and normative, $b = 0.77$, $p < 0.001$ and distinctive accuracy, $b = 0.14$, $p < 0.001$ of values. The levels of accuracy were similar to those expected after accounting for the nested structure, thus the results from the entire sample are presented in the Results section.

2.4.2. Reliability of visibility ratings

A total of 89 participants rated visibility dimensions for the value items with an average of 30 raters per item; 158 participants rated trait items with an average of 32 raters per item. Inter-rater reliability was examined by correlating each individual set of

³ Additional measures not used in the current analyses included the Miller Social Intimacy Scale (Miller and Lefcourt, 1982) and the Satisfaction With Life Scale (Diener, Emmons, Larsen, & Griffin, 1985).

Table 1
Intercorrelations among visibility dimensions and accuracy for personal values.

	1	2	3	4	5	6	7	8	9	10
1. Self-other agreement	–									
2. Consensus	0.70	–								
3. Total visibility	–0.03	–0.12	–							
4. Ease of imagining confirming behaviors	0.05	–0.06	0.88	–						
5. Ease of imagining disconfirming behaviors	0.09	0.01	0.88	0.77	–					
6. Frequency of occasions for confirming behaviors	–0.08	–0.12	0.89	0.78	0.71	–				
7. Frequency of occasions for disconfirming behaviors	–0.10	–0.05	0.87	0.68	0.89	0.83	–			
8. Easiness to judge in others	0.12	0.01	0.89	0.89	0.80	0.84	0.76	–		
9. Easiness to judge in yourself	–0.12	–0.22	0.94	0.78	0.78	0.83	0.81	0.85	–	
10. Favorability	–0.10	–0.20	0.69	0.46	0.47	0.47	0.41	0.38	0.62	–

Note: N = 56. Bold values indicate $p < .05$.

Table 2
Intercorrelations among visibility dimensions and accuracy for personality traits.

	1	2	3	4	5	6	7	8	9	10
1. Self-other agreement	–									
2. Consensus	0.45	–								
3. Total visibility	–0.08	0.14	–							
4. Ease of imagining confirming behaviors	0.05	0.17	0.92	–						
5. Ease of imagining disconfirming behaviors	–0.03	0.25	0.86	0.86	–					
6. Frequency of occasions for confirming behaviors	0.01	0.12	0.92	0.91	0.81	–				
7. Frequency of occasions for disconfirming behaviors	–0.07	0.12	0.89	0.83	0.86	0.92	–			
8. Easiness to judge in others	0.08	0.22	0.94	0.94	0.86	0.94	0.87	–		
9. Easiness to judge in yourself	–0.03	0.17	0.89	0.80	0.84	0.82	0.83	0.86	–	
10. Favorability	– 0.29	–0.10	0.42	0.18	0.05	0.16	0.10	0.17	0.17	–

Note: N = 100. Bold values indicate $p < .05$. Favorability scores of reverse coded traits were reverse coded.

ratings with all the other raters' sets of ratings, across all items. Ratings provided from three raters for values were negatively correlated with those of the other raters. Funder and Dobroth (1987) removed insufficient raters from their analyses, but this did not significantly impact reliability. Therefore, in the current study, ratings provided by all of the raters were used. Using Cronbach's alpha to measure reliability, the average reliability for values was 0.85 and the average reliability for traits was 0.94.

2.4.3. Computing visibility of values and traits

Composite visibility scores were computed by first averaging the ratings for each dimension on each value and trait item. All six dimension composites were averaged together for each value and trait item, yielding one total visibility score for each item. See Tables 1 and 2 for intercorrelations among the visibility dimensions.

2.4.4. Computation of item-level accuracy scores for values and traits

Consistent with the analyses performed by Funder and Dobroth (1987), self-other agreement for value and trait items was determined by: (1) averaging the two acquaintance ratings of each target for each item, and (2) correlating the mean acquaintance ratings with self-ratings for each trait and value item to determine self-other agreement. Each participant rated two other acquaintances for a total of 408 acquaintance ratings. Using intraclass correlations between each of the acquaintance pairs, the average reliability for values was 0.42 and the average reliability for traits was 0.65. Cronbach's alphas for the entire sample were 0.98 and 0.99, respectively.

The accuracy correlations were at the item-level, meaning that a correlation was calculated for each item between self-ratings and the average acquaintance ratings, across all judge-target pairs. Consensus was calculated by correlating the ratings of the two acquaintances of the same target, for each value and trait item. This resulted in 100 correlations for traits and 56 correlations for values for each outcome: self-other agreement and consensus.

When considering accuracy of personality judgments, the correlation does not constitute a strictly dichotomous judgment, but a relative standing on continuous scale. The accuracy scores were used to allow self-other agreement and consensus to be correlated with visibility and favorability ratings of the items.

2.4.5. Computation of profile accuracy scores for values and traits

Profile correlations for self-other agreement were computed between the self and the average of the two acquaintances across all of the items for each target. Profile correlations for consensus were computed by correlating the scores of the two acquaintance ratings across all the items for each target. These correlations were used to generate a score for each target so that accuracy for traits and values could be compared.

3. Results

3.1. Ratings of values and traits

To determine whether values or traits are more visible (Hypothesis 1), averages were computed across all raters for each value and trait item. An independent samples t -test showed that values ($M = 6.53$, 95% CI [6.36, 6.70]⁴, $SD = 0.66$) were rated as more visible than traits ($M = 6.15$ [6.01, 6.29], $SD = 0.89$), $t(154) = 2.81$, $p = 0.01$, $d = 0.48$. Desirability of each value and trait item was assessed with the question, "How favorably or unfavorably would you regard a person who possessed this trait (or value)?" Half of the trait items were negatively valenced, therefore the respective favorability ratings for those items were reverse scored and then aggregated with the ratings of the positively valenced items to reflect an overall favorability score for traits. An independent samples t -test revealed that favorability ratings of values ($M = 7.08$ [6.78, 7.38], $SD = 1.13$) were significantly greater than favorability ratings of traits ($M = 6.03$ [5.75, 6.31], $SD = 1.76$), $t(154) = 4.01$, $p < 0.001$, $d = 0.70$.

⁴ Values in brackets represent 95% confidence intervals.

Frequency of the value and trait items were assessed with two questions “In the course of normal social interaction, how frequently do occasions arise that would allow for behaviors that confirm [disconfirm] this trait [value]?” An independent samples *t*-test revealed that ratings of frequency of occasions to confirm values ($M = 6.72$ [6.61, 6.83], $SD = 0.73$) were significantly greater than that of traits ($M = 6.29$ CI [6.12, 6.46], $SD = 1.08$), $t(154) = 2.63$, $p = 0.01$, $d = 0.47$. However, ratings of occasions to disconfirm values ($M = 6.03$ [5.92, 6.14], $SD = 0.69$) were not significantly different than that of traits ($M = 5.90$ [5.75, 6.05], $SD = 0.93$), $t(154) = 0.89$, $p = 0.38$, $d = 0.16$.

Regression analyses for values and for traits revealed that favorability predicted frequency for values, $b = 0.30$, $p = 0.001$, but not for traits, $b = 0.10$, $p = 0.10$. Furthermore, multiple regressions revealed that both favorability and frequency predicted overall visibility for values and traits, with frequency as the stronger predictor of overall visibility for both values, $b = 0.67$, $p < 0.001$; $b = 0.20$, $p < 0.001$, and traits, $b = 0.72$, $p < 0.001$; $b = 0.14$, $p < 0.001$.

3.2. Comparison of accuracy for judging values and traits

To test Hypothesis 2, that values would be judged more accurately than traits, accuracy was computed both at the item level and at the profile level. Analyses were also performed for trait and value domains at the scale level. (See [Table S2 in the online supplemental materials.](#))

3.2.1. Item-level accuracy

Item-level accuracy allows for the overall comparison of values and traits across all of the judges, which addresses the question of whether accuracy differs between judgments of values and traits. Item-level accuracy is consistent with the analyses in [Funder and Dobroth \(1987\)](#).

An independent samples *t*-test based was used to compare the accuracy of items used to assess traits and items used to assess values. Self-other agreement was higher for trait items ($M = 0.22$ [0.20, 0.24], $SD = 0.12$) than value items ($M = 0.14$ [0.12, 0.16], $SD = 0.11$), $t(154) = 3.81$, $p < 0.001$, $d = 0.69$. Consensus was also higher for trait items ($M = 0.18$ [0.17, 0.19], $SD = 0.09$) than value items ($M = 0.10$ [0.08, 0.12], $SD = 0.12$), $t(154) = 4.98$, $p < 0.001$, $d = 0.75$.

To determine whether differences in visibility affected these differences in accuracy scores, a multiple regression examining self-other agreement between values and traits was conducted, while holding visibility constant. Visibility was entered into the model in step 1 and a dummy-coded variable representing the category of trait or value was entered in step 2. The dependent variable was self-other agreement. The analysis revealed that even while controlling for visibility, traits were judged more accurately than values, $\beta = 0.07$, $p < 0.001$.

Another multiple regression examining consensus between values and traits, while holding visibility constant, was conducted and revealed traits to be judged with higher levels of consensus than values, $\beta = 0.09$, $p < 0.001$.

3.2.2. Overall profile accuracy

Profile accuracy allows direct comparisons between judges for accuracy of value ratings and accuracy of trait ratings, which addresses the question of whether accuracy of values and trait ratings differ between judges. Profile-level accuracy is consistent with the analyses in [Dobewall et al. \(2014\)](#).

The resulting correlations were *z* transformed and a paired-samples *t*-test was conducted to compare the levels of accuracy that each participant obtained for judgments of traits and values. Although both traits and values were judged with significant levels of accuracy, traits ($M = 0.55$ [0.52, 0.57]), $SD = 0.20$) were judged

more accurately than values ($M = 0.42$ [0.39, 0.44], $SD = 0.18$), $t(203) = 8.50$, $p < 0.001$, $d = 0.73$.

Consensus between the two acquaintances, computed using profile correlations, was also higher for traits ($M = 0.38$ [0.35, 0.41], $SD = 0.22$) than for values ($M = 0.30$ [0.27, 0.33], $SD = 0.22$), $t(203) = 4.27$, $p < 0.001$, $d = 0.36$.

3.2.3. Distinctive profile accuracy

The influence of normativeness, or what people are like on average, can be an important piece of information in some personality research, however it can be problematic when assessing accuracy with profile correlations. When correlating the profiles of two people, the resulting correlation may not reflect how similar the specific ratings of two people are to one another, but instead reflect two people's ratings of the “average person.” [Cronbach \(1955\)](#) criticized profile correlations as assessments of accuracy, because similarity between the ratings of two people could actually be due to the raters' ability to rate the average person instead of the ability to rate a specific person. Normativeness of profile correlations may artificially inflate overall accuracy because what is being assessed is the degree of similarity between two “average people”. The issue can be addressed by calculating a distinctive profile by subtracting the normative profile (or the average of the items across persons) from the individual profiles, which then represents the unique aspects of each set of ratings ([Furr, 2008](#)).

Accuracy correlations computed with the distinctive profile for the accuracy criteria reflect the consistency between the judgment and the unique aspects of the accuracy criterion profile. Using this method, distinctive accuracy for traits and values was calculated by correlating the composite peer-judgments of items with the distinctive self-rating profile scores for that target for all value and trait items. Distinctive accuracy scores for values and traits were both *z* transformed and compared with a paired-samples *t*-test. Distinctive accuracy for traits ($M = 0.17$ [0.13, 0.21], $SD = 0.28$) was not significantly different from distinctive accuracy for values ($M = 0.14$ [0.11, 0.17], $SD = 0.23$), $t(203) = -1.5$, $p = 0.13$, $d = 0.12$, which is not consistent with the finding for overall accuracy.

3.3. Visibility and accuracy

To test Hypothesis 3, that more visible values/traits would be judged more accurately than less visible values/traits, the composite visibility scores for each value and trait item were correlated with item-level self-other agreement and consensus scores. Relationships were not found between the visibility scores and any of the accuracy scores (self-other agreement for values $r = -0.03$ [-0.29, 0.23], $p = 0.81$, and traits $r = -0.08$ [-0.27, 0.12], $p = 0.45$; consensus for values, $r = -0.12$ [-0.37, 0.15], $p = 0.39$, and traits, $r = 0.14$ [-0.06, 0.33], $p = 0.17$. Although there is not a test to assess whether these correlations differ significantly, confidence intervals for each of the correlations indicate some overlap between values and traits. (See [Figs. S1 and S2 in the online supplemental materials.](#))

An additional analysis examined the prediction of self-other agreement by visibility and type of characteristic (trait or value). In a multiple regression, traits and values were dummy coded as 0 and 1, respectively, and were included as predictors along with visibility, while self-other agreement was the outcome. The type of characteristic predicted self-other agreement, $b = -0.07$, $p < 0.001$, while visibility did not, $b = -0.01$, $p = 0.43$.

To determine if separate aspects of the visibility scores were related to accuracy, each of the individual visibility dimensions were correlated with self-other agreement and consensus. Significant associations were found for traits between self-other agreement and the favorability dimension, $r = -0.29$ [0.10, 0.46] $p = 0.004$, and consensus and the ease of judgment of others

dimension, $r = 0.22$ [0.02, 0.40], $p = 0.03$. The other dimensions did not yield significant relationships with any type of accuracy for traits or values (see Tables 1 and 2).

Furthermore, an exploratory question was to examine which was the best predictor of accuracy, visibility or favorability. A multiple regression with favorability and overall visibility as predictors and self-other agreement as the outcome revealed that only favorability predicted self-other agreement for traits, $b = -0.02$, $p = 0.003$ but not for values, $b = -0.015$, $p = 0.42$.

3.3.1. Differences between most and least visible items

To determine if there were possible differences between levels of self-other agreement for the most visible and least visible values and traits (H3), confidence intervals were calculated for the correlation coefficients. The correlation between visibility and self-other agreement for the most visible value type (conformity $r = 0.19$ [0.05, 0.32]), and the least visible value type (universalism $r = 0.16$ [0.02, 0.29]) have highly overlapping confidence intervals. The correlation between visibility and self-other agreement for the most visible trait domain (extraversion $r = 0.22$ [-0.25, 0.60], $p = 0.35$) and the least visible trait domain (openness $r = -0.32$ [-0.67, 0.14], $p = 0.18$) also have overlapping confidence intervals. (See Tables S3 and S4 for the most and least visible value and trait items in the online supplemental materials.)

3.3.2. Accuracy and visibility for individual value types and the Big Five personality domains

To test the portion of Hypothesis 3, that the stimulation and tradition value types would be rated the most visible and judged the most accurately, item-level self-other agreement correlations were computed for the higher-order factors of self-transcendence, self-enhancement, openness to change, and conservation, and for individual value types, by averaging the z transformed item correlations for items that assess each value type. Item-level self-other agreement correlations were also computed for each of the Big Five personality traits in the same way. Table 3 shows the self-other agreement correlations for the value types, higher-order values,

and the Big Five domains, ordered from highest to lowest agreement. Extraversion and neuroticism had the highest self-other agreement for personality traits, and stimulation and tradition had the highest self-other agreement for value types, which partially confirms hypothesis 3. The findings of the greatest level of self-other agreement for the higher-order values of Conservation (0.19) and Self-transcendence (0.16) were partially consistent with those of Dobewall et al. (2014).

Means of the composite visibility scores of the 10 value types ranged from 6.05 for universalism to 7.27 for conformity. The findings do not support Hypothesis 3, as stimulation and tradition were not rated as the most visible value types. The means for the Big Five personality domains range from 5.18 for openness to experience to 6.76 for extraversion (see Table 3).

3.3.3. Controlling for acquaintanceship

Previous research has suggested that the relationship between visibility and accuracy that is demonstrated in low and moderately acquainted dyads disappears in highly acquainted dyads (Paunonen, 1989). It is possible that the current null findings between visibility and accuracy are due to the highly acquainted nature of the current sample. To test for the influence of acquaintanceship level, the sample was divided into high and low acquaintance groups by a median split of relationship length (Mdn = 2 years). Item-level self-other agreement and consensus were determined for the low acquaintance group and for the high-acquaintance groups, for each value and trait item. Self-other agreement was then correlated with visibility scores, across items. There were no significant associations between visibility and self-other agreement for values for either low acquaintance, $r = 0.07$, $p = 0.64$, or high acquaintance, $r = -0.07$, $p = 0.59$, or for traits for either low acquaintance, $r = -0.10$, $p = 0.32$, or high acquaintance, $r = 0.07$, $p = 0.47$. Consensus for both the high and low acquaintance groups was also correlated with visibility scores, which again revealed no significant associations for values for either low acquaintance, $r = -0.18$, $p = 0.19$, or high acquaintance, $r = -0.03$, $p = 0.81$, or for traits for either low acquaintance, $r = 0.14$, $p = 0.17$, or high acquaintance, $r = 0.12$, $p = 0.23$.

Table 3

Visibility and favorability composite means, item-level self-other agreement and consensus of personality domains (NEO-PI-R), higher-order values, and value types (SVS).

NEO	Visibility M	Desirability M	Self-other agreement r	Consensus r
Extraversion	6.76	6.40	0.33***	0.25**
Neuroticism	5.86	3.34	0.25**	0.17*
Conscientiousness	6.29	6.41	0.10	0.18*
Agreeableness	6.44	7.35	0.19*	0.15
Openness	5.18	6.04	0.24**	0.16*
SVS				
Conservation	6.59	7.04	0.19*	0.08
Conformity	7.27	8.14	0.19*	0.07
Security	6.45	7.03	0.17*	0.06
Tradition	6.06	6.31	0.22**	0.12
Self-enhancement	6.73	6.52	0.08	0.05
Achievement	6.95	7.67	0.06	0.01
Power	6.50	5.32	0.11	0.11
Self-transcendence	6.48	7.50	0.16*	0.13
Benevolence	6.90	7.92	0.16*	0.11
Universalism	6.05	7.08	0.16*	0.15
Openness-to-change	6.44	7.20	0.12	0.14
Self-direction	6.46	7.40	0.08	0.10
Stimulation	6.41	6.87	0.20**	0.20**
Hedonism	6.43	6.68	0.08	-0.01

Note. Visibility and favorability ratings: N = 247.

Self-other agreement and consensus: N = 204.

* $p < 0.05$.

** $p < 0.01$.

*** $p < 0.001$.

4. Discussion

We predicted that values would be rated as more visible than traits (H1), which was supported in the current study. We also predicted that increased visibility of values would be due to greater availability and favorability of values, which was supported.

For Hypothesis 2, we predicted that personal values would be judged with greater accuracy than personality traits, while visibility was held constant. Instead, the opposite was found at both the item and profile levels, such that traits were judged significantly more accurately than values. The finding was surprising due to the result that values were rated as both more favorable and visible than traits. Some personal values can be judged accurately, but as a whole, values are judged less accurately than traits. However after removing the normative profile from the ratings, accuracy of traits was no longer significantly different from accuracy of values.

We also predicted that, similar to previous findings for traits, more visible values would be judged more accurately than less visible values (H3), which was not supported by the data. Although some value items rated high in visibility were also judged accurately, in general the most visible values were not judged the most accurately. Furthermore, the most visible traits were also not judged the most accurately. In addition, overlapping confidence intervals suggest that the relationship between visibility and accuracy for values may not be significantly different from the relationship between visibility and accuracy for traits. Overall, visibility and accuracy were not significantly related in this sample. Furthermore examining the relationship between accuracy and each of the individual visibility dimensions did not yield significant results, except for the favorability dimension which predicted accuracy of traits. However, contrary to expectations, favorability was not related to accuracy for values.

We also expected that the stimulation and tradition value types to be the most visible and judged more accurately than the other value types (H3). The prediction was partially supported in this study, as the most visible value types were actually conformity, achievement, and benevolence, and although not rated as the most visible, the two value types of stimulation and tradition were judged the most accurately in terms of self-other agreement at the item-level.

Examining visibility and accuracy in different levels of acquaintanceship did not make a difference in terms of the non-association between the two variables. It appears that this sample was so highly acquainted that visibility of traits or values did not have a relationship with accuracy.

4.1. Relations with previous research

Interestingly, the results are inconsistent with those of [Funder and Dobroth \(1987\)](#) who found subjective visibility and accuracy of traits to be moderately correlated and that the most visible trait items were also those with the highest accuracy. It is curious that the current results did not replicate those in the former study, given that most controversies over observability and self-other agreement typically stem from the problem of trait extremes, which is to say that the bi-polar ends of the scales of traits are usually more visible and also judged more accurately than the middle of the trait scales. The visibility measure used by [Funder and Dobroth \(1987\)](#) and [Rothbart and Park \(1986\)](#) took into account both the confirmability and disconfirmability of traits. It was assumed in the current research that using the validated measure would eliminate any problems that often arise when assessing visibility of traits. The problem with trait extremes most likely is not an issue in this study, and does not explain the null findings.

The current results are also inconsistent with those of [John and Robins \(1993\)](#), who assessed visibility (which they called observability) of the Big Five traits with one question, and correlated the scores with both self-other agreement and peer agreement. Observability was found to be correlated with both self-other agreement and peer-agreement and, like [Funder and Dobroth \(1987\)](#), indicated that more observable traits were judged more accurately, which again is not what the current study revealed. However, consistent with [John and Robins \(1993\)](#) and [Vazire \(2010\)](#) regarding the evaluativeness of traits, favorability in the current study had a significant negative relation with self-other agreement. The negative association between favorability and self-other agreement was not present for consensus, suggesting that people may rate themselves higher on more favorable traits than do their acquaintances.

An explanation for the differences in results from earlier findings is that previous studies examining visibility and accuracy were conducted with college students in the late 1980's and early 1990's. The cohort of undergraduates today may differ in the way they judge one another based on the advent of social media, reality TV shows, and the like. Judging the motives and personalities of others one does not know or does not know well appears to have turned into a national pastime. It may be that when thinking about visibility of traits and values, participants thought of the average person they know in a distant way, which is different from those they know well and interact with every day.

[Paunonen \(1989\)](#) found that among highly acquainted individuals, visibility of traits is not necessarily related to accuracy of trait judgment, which the current findings support. The method assessing familiarity between acquaintances in the current study was the same as that used by [Paunonen \(1989\)](#), which rated acquaintanceship on a scale of 1 (*low*) to 9 (*high*). He found that at high levels of acquaintanceship (above a score of 5) visibility was no longer related to accuracy of trait judgment. Acquaintanceship for the current study had a mean score of 7.29, indicating that most participants were highly acquainted. At such high-levels of acquaintanceship, the visibility of cues may no longer influence accuracy, because judges can detect both high and low visibility cues due to the amount of time previously spent with the target in multiple contexts ([Paunonen \(1989\)](#)). In support of Paunonen's argument, the dyads in the current sample were highly acquainted to the degree that visibility was not related to accuracy. One implication of the results may be that a relationship of 6 months is an adequate length to establish a level of high of acquaintanceship for future studies.

Other current findings consistent with previous research are that stimulation and tradition value types are judged the most accurately. [Pozzebon and Ashton \(2009\)](#) found that self-reported stimulation and tradition value types had the most peer-reported value type behaviors associated with them. However, the visibility ratings do not support their findings linking value ratings and peer-reported behaviors, as the most visible value types (e.g., conformity, achievement, and benevolence) were some of the least predictive value types of peer-reported behaviors in the previous study. Encouragingly, some values are more accurately judged (tradition, stimulation, and conformity) than others (self-direction and achievement), similar to how some traits (extraversion) are more accurately judged than others (openness). It appears that individuals can accurately judge some values of well-known acquaintances, nevertheless, the accuracy of value judgments does not seem to be related to visibility or favorability of values.

Consistent with the research of [Dobewall et al. \(2014\)](#), the current findings reflect that values can be accurately judged, but traits are judged slightly more accurately. Our results show the same superiority of traits over values but to an even larger degree. The results of the current study and [Dobewall et al. \(2014\)](#) are both

consistent with other research that found personality traits to be more predictive of behavior than values (Bardi & Schwartz, 2003) and even more predictive of some value-relevant behaviors (Pozzebon & Ashton, 2009).

4.2. Implications

4.2.1. Theoretical

Operating under the framework of the Realistic Accuracy Model (RAM), it appears that in the current sample, participants appraised values as being visible, frequent, and favorable which is consistent with the availability stage of RAM. However, the participants actually making the personality judgments may not have perceived the particular cues relevant to the judgment of values as easily detected as was anticipated based on the visibility ratings. Furthermore, the cues relevant to personal values may not have been utilized correctly.

Inversely, traits were rated as being less visible, frequent, and favorable, suggesting they would have fewer available cues, but were in fact judged more accurately. The availability stage was therefore most likely not responsible for accuracy for these individuals, but due to the high acquaintanceship level of the participants, the cues that were available were detected and utilized correctly. Regarding favorability of items, when examined independently, favorability ratings predicted lower self-other agreement for traits, which may mean that people may be better at detecting cues for less favorable traits than more favorable traits. Another interpretation is that others are better than the self at utilizing cues regarding favorable traits. Future research could examine the differences between traits and values at the detection and utilization stages. Another idea is to explore if accurate judgments of values can be made between people of zero acquaintance, and if there is a relationship between visibility of values with accuracy of judgments in that context.

In consideration of moderators, there may have been a difference in the abilities of the judges and the judgability of the targets, but most importantly there was a difference in the quality of constructs. Regarding the good trait moderator of RAM, personality traits are indeed better *constructs* than values when making judgments because traits can be judged more accurately.

4.2.2. Applied

The findings in the current study have potential to be applied in many settings. The belief that some traits or values are more visible than others may generalize to individuals in professional, academic, legal, and even personal settings. People may imagine they are competent at accurately judging values in those settings, but in reality the values of others may be elusive. There are implications for job interviews, where an interviewer believes he or she can accurately judge what motivates and guides an individual looking for a job. The findings may also be applied to those working in human resources or schools, who believe they know how to motivate individuals. If there is a disconnect between what supervisors believe motivates individuals and what truly motivates their employees, the goals of the organization may not be met. Finally the findings are especially applicable in legal settings where jury selection is crucial. The values of potential jurors are important to lawyers representing a client and if lawyers believe values are more visible than they really are, decisions made about jurors via interviews could impact the outcome of trials in undesirable ways. Measuring the values of individuals in the aforementioned settings in a more systematic way may be called for. The Schwartz Value Survey could be utilized by businesses, schools, and courts in a battery of other personality tests. Combining the use of surveys and interviews may yield more accurate judgments and better outcomes in these settings.

In the personal domain, the findings imply that not only do people believe the values of others are more visible than they really are, but likewise their own values are more visible to others than they really are. The implication mirrors the social psychological concept of the illusion of transparency (Gilovich, Savitsky, & Medvec, 1998), where individuals feel that their inner states are more observable to others than they actually are. The premise of the illusion of transparency is that people are “anchored” to their own experience when making judgments of perspectives of others, and they have difficulty adjusting appropriately from that anchor. In the current study, it may be that participants believed that their own values were more visible to others, which they interpreted to mean that the values of others would likewise be more visible to them.

4.3. Limitations

One limitation is that the visibility ratings were collected online, which has pros and cons. For example, online measures are convenient for participants, but there is no way of controlling the administration of the measures or ensuring that people have minimal distractions.

A second limitation regarding accuracy ratings is that the 100 trait items were not administered randomly, but instead were listed in order by trait domain (e.g. all 20 of the items for extraversion were listed in order together) which could influence how the raters responded to the questions. It is possible that participants could see a pattern in the questions and perceive what trait was being assessed, which inevitably could lead to socially desirable responding.

A final limitation is that the sample was highly acquainted which may account for the non-replication of previous work regarding visibility and accuracy. Future research could examine the relation between visibility and accuracy of personal values in unacquainted dyads.

4.4. Future research

An additional idea for future research would be to apply the Self-Other Knowledge Asymmetry model (SOKA; Vazire, 2010) to values. The SOKA model posits that some aspects of personality are judged more accurately by the self and other aspects are judged more accurately by other people. In particular, traits low in observability and low in evaluativeness, such as neuroticism, are better judged by the self; and traits high in observability but low in evaluativeness, such as extraversion, are judged better by others. It would be interesting to observe if this model also applies to personal value types. Some less observable, less evaluative values may be best judged by the self and more observable, less evaluative, values may be best judged by others. For example, value items that are more observable, such as *wisdom* and *helpful*, might be judged best by others. Some support from this idea comes from the current finding that consensus was higher for those value items than was self-other agreement. Values such as *family-security*, *humble*, and *honest* were also rated as highly visible, but self-other agreement was higher for those values and not significant for consensus. Values that are less observable but more favorable may be judged better by the self.

It would also be beneficial to the field of accuracy of personality judgment to examine the values of those who are more or less accurate judges to gain an understanding of some of the motives held by “good judges.” Due to a potentially greater ability to detect and utilize cues, those who hold self-transcendent values may be more accurate at judging the traits and values of others than those who hold self-enhancement values.

Another suggestion for future research would be to examine accuracy and visibility of personal values with the level of initial acquaintanceship as a moderator. Previous research has found increases in self-other agreement of personality traits with increasing interaction length between individuals who had just met (Blackman & Funder, 1998). A future study could determine if accuracy of personal values would similarly improve depending on the amount of acquaintanceship.

4.5. Conclusion

The current study found that traits are judged more accurately than values in a sample of well-acquainted individuals, while values in general are rated as more visible than traits. It may be that this group of participants believes that people will be good judges of the motives behind the behaviors of others, while in reality they are not, even for people whom they feel they know well. Inversely, participants did not perceive that traits can be judged as easily as values, but in reality others judge traits fairly accurately as compared to values. There appears to be a disconnect between how easy people think it will be to judge others and how they actually judge their acquaintances. Future directions in research of accuracy of personality judgment can explore why this disconnect exists and perhaps discover how to narrow the gap between people's perceptions of judgment and actual judgment of the personalities of others.

Appendix A

A.1. Ratings of trait and value items

A.1.1. Visibility

1. *Imaginability of confirming behaviors*: "How easy is it to imagine specific, observable behaviors that would provide confirmation of that trait(value)?"
2. *Imaginability of disconfirming behaviors*: "How easy is it to imagine specific, observable behaviors that would disconfirm (provide evidence against) that trait(value)?"
3. *Frequency of occasions allowing confirming behaviors*: "In the course of normal social interaction, how frequently do occasions arise that would allow for behaviors that confirm this trait(value)?"
4. *Frequency of occasions allowing disconfirming behaviors*: "In the course of normal social interaction, how frequently do occasions arise that would allow for behaviors that disconfirm this trait(value)?"
5. *Easiness*: "How difficult or easy would it be to judge the degree to which another person had this trait?"
6. *Easiness for others*: "How difficult or easy would it be for others to judge the degree to which you had this trait?"

A.1.2. Favorability

7. *Favorability*: "How favorably or unfavorably would you regard a person who possessed this trait?"

Appendix B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.jrp.2016.10.009>.

References

Bardi, A., & Schwartz, S. H. (2003). Values and behavior: Strength and structure of relations. *Personality and Social Psychology Bulletin*, 29, 1207–1220. <http://dx.doi.org/10.1177/0146167203254602>.

- Beer, A., & Watson, D. (2008). Personality judgment at zero acquaintance: Agreement, assumed similarity, and implicit simplicity. *Journal of Personality Assessment*, 90, 250–260.
- Biesanz, J. C. (2010). The social accuracy model of interpersonal perception: Assessing individual differences in perceptive and expressive accuracy. *Multivariate Behavioral Research*, 45(5), 853–885.
- Bilsky, W., & Schwartz, S. H. (1994). Values and personality. *European Journal of Personality*, 8, 163–181.
- Blackman, M. C., & Funder, D. C. (1998). The effect of information on consensus and accuracy in personality judgment. *Journal of Experimental Social Psychology*, 34, 164–181.
- Colvin, C. R. (1993). "Judgable" people: Personality, behavior, and competing explanations. *Journal of Personality and Social Psychology*, 64, 861–873.
- Costa, P. T., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology*, 54, 853–863. <http://dx.doi.org/10.1037/0022-3514.54.5.853>.
- 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.
- Dobewall, H., Aavik, T., Konstabel, K., Schwartz, S. H., & Realo, A. (2014). A comparison of self-other agreement in personal values versus the Big Five personality traits. *Journal of Research in Personality*, 50, 1–10. <http://dx.doi.org/10.1016/j.jrp.2014.01.004>.
- Fleeson, W., Furr, R. M., Jayawickreme, E., Meindl, P., & Helzer, E. G. (2014). Character: The prospects for a personality-based perspective on morality. *Social and Personality Psychology Compass*, 8, 178–191. <http://dx.doi.org/10.1111/spc3.12094>.
- Funder, D. C. (1991). Global traits: A neo-Allportian approach to personality. *Psychological Science*, 2, 31–39.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102, 652–670. <http://dx.doi.org/10.1037/0033-295X.102.4.652>.
- Funder, D. C., & Drobth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology*, 52, 409–418. <http://dx.doi.org/10.1037/0022-3514.52.2.409>.
- Furr, R. M. (2008). A framework for profile similarity: Integrating similarity, normativeness, and distinctiveness. *Journal of Personality*, 76, 1267–1316. <http://dx.doi.org/10.1111/j.1467-6494.2008.00521.x>.
- Gilovich, T., Savitsky, K., & Medvec, V. H. (1998). The illusion of transparency: Biased assessments of others' ability to read one's emotional states. *Journal of Personality and Social Psychology*, 75, 332–346.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48, 26–34. <http://dx.doi.org/10.1037/0003-066X.48.1.26>.
- Human, L. J., & Biesanz, J. C. (2013). Targeting the good target: An integrative review of the characteristics and consequences of being accurately perceived. *Personality and Social Psychology Review*, 17, 248–272. <http://dx.doi.org/10.1177/1088868313495593>.
- IPIP NEO-PI domains. International personality item pool: A scientific laboratory for the development of advanced measures of personality traits and other individual differences (<<http://ipip.ori.org/>>). Internet Web Site.
- 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(4), 521–551.
- Letzring, T. D. (2008). The good judge of personality: Characteristics, behaviors, and observer accuracy. *Journal of Research in Personality*, 42, 914–932. <http://dx.doi.org/10.1016/j.jrp.2007.12.003>.
- Letzring, T. D., & Human, L. J. (2013). An examination of information quality as a moderator of accurate personality judgment. *Journal of Personality*. <http://dx.doi.org/10.1111/jopy.12075>.
- McAdams, D. P. (1995). What do we know when we know a person? *Journal of Personality*, 63(3), 365–369. <http://dx.doi.org/10.1111/1467-6494.ep9510042296>.
- McCrae, R. R., & Costa, P. T. (1990). *Personality in adulthood*. New York: Guilford.
- McCrae, R. R., & John, O. J. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175–215.
- Miller, R. S., & Lefcourt, H. M. (1982). The assessment of social intimacy. *Journal of Personality Assessment*, 46, 514–518.
- Nicholson, I. A. M. (2003). *Inventing personality*. Washington, DC: American Psychological Association.
- Paunonen, S. V. (1989). Consensus in personality judgments: Moderating effects of target-rater acquaintanceship and behavior observability. *Journal of Personality and Social Psychology*, 56, 823–833. <http://dx.doi.org/10.1037/0022-3514.56.5.823>.
- Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry*, 5(2), 103–113. http://dx.doi.org/10.1207/s15327965pli0502_1.
- Pozzobon, J. A., & Ashton, M. C. (2009). Personality and values as predictors of self- and peer-reported behavior. *Journal of Individual Differences*, 30, 122–129. <http://dx.doi.org/10.1027/1614-0001.30.3.122>.
- Roccas, S., Sagiv, L., Schwartz, S. H., & Knafo, A. (2002). The big five personality factors and personal values. *Personality and Social Psychology Bulletin*, 28, 789–801. <http://dx.doi.org/10.1177/0146167202289008>.
- Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology*, 50, 131–142. <http://dx.doi.org/10.1037/0022-3514.50.1.131>.

- Schwartz, S. H. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology*, 53, 550–562.
- Schwartz, S. H., & Bilsky, W. (1990). Toward a theory of universal content and structure of values: Extensions and cross-cultural replications. *Journal of Personality and Social Psychology*, 58, 878–891. <http://dx.doi.org/10.1037/0022-3514.58.5.878>.
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in experimental social psychology*, 35 (pp. 1–65). New York: Academic Press.
- Schwartz, S. H., & Sagie, G. (2000). Value consensus and importance: A cross-national study. *Journal of Cross-Cultural Psychology*, 31, 465–497. <http://dx.doi.org/10.1177/0022022100031004003>.
- Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology*, 32, 519–542. <http://dx.doi.org/10.1177/0022022101032005001>.
- Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, 98, 281–300. <http://dx.doi.org/10.1037/a0017908>.
- Vecchione, M., Alessandri, G., Barbaranelli, C., & Caprar, G. (2011). Higher-order factors of the big five and basic values: Empirical and theoretical relations. *British Journal of Psychology*, 102, 478–498. <http://dx.doi.org/10.1111/j.2044-8295.2010.02006.x>.
- Vogt, D. S., & Colvin, C. R. (2003). Interpersonal orientation and the accuracy of personality judgments. *Journal of Personality*, 71, 267–295.