Perceptions of personality in text-based media and OSN: A meta-analysis

Konstantin O. Tskhay *, Nicholas O. Rule

University of Toronto, Canada

ARTICLE INFO

Article history:
Available online 28 December 2013

Keywords:
Big Five traits
Facebook
OSN
Person perception
Meta-analysis
Zero-acquaintance

ABSTRACT

Research has shown that personality can be accurately perceived at zero-acquaintance. Although most of this work has focused on physical appearance, a growing number of studies has suggested that personality may be perceptible from other sources of information. In the current meta-analysis, we examined studies that reported accuracy and consensus effects for the perception of the Big Five traits from text-based media and online social network websites. We found substantial consensus for all five traits. Moreover, extraversion, openness to experience, agreeableness, and conscientiousness were perceived accurately. Importantly, we provide the aggregate effect sizes that researchers might expect when examining similar phenomena to help guide future studies and discuss several potential avenues for valuable additional research in personality and person perception.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Personality is an important factor in individuals’ lives. Individual trait differences have been shown to predict academic achievement, job outcomes, general health, and success in romantic and interpersonal relationships (Caspi, Roberts, & Shiner, 1992, 1993; Connelly & Ones, 2010; Gosling, Ko, Mannarelli, & Morris, 2002). The Five-Factor Model of personality has become the dominant framework that researchers utilize in understanding the internal stability of individuals’ traits (John & Srivastava, 1999; McCrae & Costa, 1999). This dimensional framework describes five broad factors of personality: extraversion, agreeableness, conscientiousness, openness to experience, and neuroticism. Although much of the research in personality has concerned how these five traits influence various life outcomes, researchers have also begun examining perceptions of personality.

Numerous studies have demonstrated that people tend to generally agree with each other when making judgments of personality (i.e., consensus) and that these impressions are frequently accurate (Albright, Kenny, & Malloy, 1988; Borkenau & Liebler, 1992, 1993; Connelly & Ones, 2010; Gosling, Ko, Mannarelli, & Morris, 2002). Much of this research has been conducted with a focus on physical appearance (Borkenau, Brecke, Möttig, & Paeckecke, 2009; Naumann, Vazire, Rentfrow, & Gosling, 2009). Although appearance can provide a surprisingly high degree of information about a person (Tskhay & Rule, 2013; Zebrowitz, 1997), there are other channels of communication that people could rely upon to get a sense of what a person might be like (Gosling et al., 2002). For example, Gosling et al. (2002) suggested that personality could be manifested in an individual’s surroundings. In their study, naïve judges entered personal spaces (bedrooms and offices) and made judgments of the owners’ personalities. The authors suggested that the accuracy found for openness to experience, extraversion, and conscientiousness was partially driven by the use of cues within the occupants’ rooms. More specifically, they proposed that people leave a trail of their actions (behavioral residue) that is indicative of their personality and that they deliberately place markers that communicate their identity (identity claims). According to this theoretical framework, a person who is high in openness to experience might unintentionally leave snowboarding equipment in her office (behavioral residue) and deliberately place a world map on the wall to indicate to others her love of travel and adventure (identity claim).

Apart from one’s physical environment, behavioral residue and identity claims might manifest in other domains. Indeed, a series of recent studies has shown that Facebook pages (e.g., Back, Stopfer, Vazire, Gaddis, Schmukle, Egloff, & Gosling, 2010; Ivcevic & Ambady, 2012), personal websites (e.g., Marcus, Machilek, & Schütz, 2006; Vazire & Gosling, 2004), blogs (Li & Chignell, 2010; Qiu, Lin, Ramsay, & Yang, 2012), e-mail addresses (Back, Schmukle, & Egloff, 2008), stories (Küfner, Back, Nestler, & Egloff, 2010), stream of thought essays (Holleran & Mehl, 2008), and even music tastes (Rentfrow & Mehl, 2008) might be indicative of one’s personality. A common link between all of these studies is the presence of deliberate or unintentional self-expressions and, importantly, communication of some form of information to others.

* Corresponding author. Address: Department of Psychology, University of Toronto, 100 St. George St., Toronto, ON M5T 3G3, Canada.
E-mail address: konstantin.tskshay@mail.utoronto.ca (K.O. Tskhay).
Writing, for example, is a medium of communication that is enriched with cues to self and personality and serves a communicative function (Pennebaker & King, 1999). Researchers have demonstrated that people scoring on different ends of personality dimensions have a persistent, consistent, and stable pattern of written expression (Chung & Pennebaker, 2008; Gill & Oberlander, 2002; Hirsh & Peterson, 2009; Holtgraves, 2011; Lee, Kim, Seo, & Chung, 2007; Pennebaker & King, 1999; Pennebaker, Mehl, & Niederoffer, 2003). Furthermore, word-use itself is suggested to be indicative of underlying personality traits (Pennebaker & King, 1999). For example, extraverted people tend to use a greater number of positive words, whereas the use of positive words was negatively correlated with neuroticism (Pennebaker & King, 1999; Yarkoni, 2010). Thus, the combination of stability in writing and the types of words used could be indicators of underlying personality constructs, supporting increased consensus and accuracy for inferring personality traits from written language.

Indeed, researchers have begun examining this very question. Across all of the studies on this topic to date, consensus for judgments of personality from written cues has usually shown statistically significant results, suggesting that people tend to form similar impressions of personality from writing samples (e.g., Li & Chignell, 2010). In most cases, the accuracy of these judgments has been found to exceed chance guessing (e.g., Holleran & Mehl, 2008; Vazire & Gosling, 2004). There are several studies, however, that have failed to find significant accuracy effects for some traits (e.g., Ivcevic & Ambady, 2012; Kühner et al., 2010). Thus, the inference of personality from writing has provided variable results and, given the relationship between personality and writing, is a relatively novel theme in person perception.

Another popular and modern medium of self-presentation and communication is online social networks (OSN; e.g., Facebook, LinkedIn). Unlike plain text, OSN are saturated with different types of information, including information about the individuals and profile pictures that can communicate what a person might be like (Hall & Pennington, 2012). This information may be useful for forming accurate impressions of extraversion, conscientiousness, and openness (Hall, Pennington, & Lueders, 2013). Other research (e.g., Kluemper, Rosen, & Mooshholder, 2012) has found that OSN provide accurate information about all Big Five traits, however. Importantly, Facebook profiles seem to communicate actual personality, not just an idealized version of the self (Back et al., 2010). Although people use OSN to accurately communicate and express aspects of their personality to others—intentionally or not—the results in the literature regarding which traits are legible from OSN have been highly varied. The present research therefore aimed to summarize these effects across studies to bring clarity to the relationship between individuals’ self-generated media and the expression and perception of personality.

1.1. Current research

To better understand the overarching capacity to infer individuals’ personality traits from writing and OSN, we quantitatively reviewed the literature on the accuracy and consensus of judging personality traits from OSN and writing. The overall effects of consensus are important because they provide researchers with an opportunity to estimate the number of raters needed to achieve high inter-rater reliability. Furthermore, we wished to investigate how accurate people are at inferring the Big Five personality traits from writing samples and OSN profiles of unacquainted strangers, in general. Additionally, we hoped to uncover new directions of inquiry that might be apparent only from a broad perspective that considers the entire field of research as a whole.

2. Method

2.1. Procedure

We searched the online academic databases PsycInfo, Google Scholar, and Web of Knowledge using the keywords personality, Big Five (including each trait), Five-Factor Model, writing, blog, computer mediated communication, OSN (including different types; e.g., Facebook), consensus, and accuracy. The search yielded 56 empirical articles or theses, all written in English. We used the cross-reference technique (Rosenthal, 1991) to identify any additional articles that could potentially meet the inclusion criteria (see below). This did not reveal any additional articles relevant to the current work. In an attempt to address the file-drawer problem, we requested any unpublished data relevant to the current meta-analysis via the online forum of the Society for Personality and Social Psychology and by directly contacting authors who had previously published in this research area. We received multiple responses and were able to obtain either the raw data or consensus and accuracy estimates for an additional eight unpublished studies.

2.1.1. Inclusion Criteria

Only 30 studies examining person perception were included; other articles did not examine personality perception and focused primarily on (a) the analysis of the content of writing as predicted from the Big Five personality dimensions (n = 11), (b) behavioral patterns in relationship to the Big Five (n = 6), (c) traits other than those described in the Five-Factor Model (n = 3), (d) measures that did not evaluate consensus and accuracy (n = 9), or (e) other reasons (e.g., case studies; n = 4). To meet inclusion, each study had to have at least two samples: raters and targets. The targets had to produce written text or have an OSN profile. They also had to complete one of the measures that estimated at least one of the Five-Factor dimensions of personality. The raters needed to have examined the materials produced by the targets and to have made a judgment of personality using a scale that assessed the same dimensions of personality. Furthermore, the studies had to provide measures of consensus and accuracy. Consensus was typically reported in terms of inter-judge agreement, measured using one of the common indices of reliability (e.g., r, ICC, Kendall’s τ, Cronbach’s x). The consensus estimate had to be convertible to an effect size estimated by the Pearson’s product-moment correlation coefficient r for agreement of a single judge with other judges in the sample. Thus, Kendall’s τs were converted to Pearson’s rs using the formulas provided by Walker (2003) and Cronbach’s x̄s were converted to single judge correlation coefficient estimates. We used the effect size r as an estimate of the correlation coefficient for a single judge. To estimate accuracy, we examined the aggregate correlations with the respective criterion (e.g., self-reported score on the personality measure). Some studies reported informant ratings of personality in addition to the self-reported personality ratings; thus, we collapsed across the two measures to form a composite. We included these studies because using both informant- and self-report is likely a better measure of the true underlying criterion, as it contains multiple indicators (Vazire, 2006). Similar to the consensus estimates, we converted Kendall’s τs to Pearson’s rs. All effect sizes were converted to Fisher’s Z prior to analyses. If authors reported finding results that did not reach conventional levels of statistical significance (z = .05) without reporting the exact estimate, we assigned an r-value of zero to provide a conservative estimate of the actual effect (n = 3).

To allow for inclusion of a greater number of studies, we did not limit inclusion based on geographic location or time period, nor did we restrict the criteria to any particular sample of targets or raters.
(although no article reported samples from clinical populations). In sum, 20 research summaries were included in the current meta-analysis. Although this number is rather conservative, these articles provided 286 independent effect sizes for accuracy and consensus across the Big Five personality traits. Sources marked with an asterisk in the References section below were included in the current analyses.

2.1.2. Coding procedure

We coded several variables in the current meta-analysis, the most important of which were the estimated effect sizes and sample sizes for each personality trait, including estimates for both accuracy and consensus. Half of the 286 effect sizes estimated consensus whereas the other half estimated accuracy. In total, the studies examined 1268 judges' ratings of 3599 targets. We also coded one moderator that could be relevant to self-expression: whether the researchers examined either written material produced by the target or the target’s OSN profile. Because it is typical for OSN to contain the owners’ pictures (Hall and Pennington, 2012) and previous research suggested that people can reliably extract information about personality from appearance (e.g., Albright et al., 1988), we decided to explore potential differences in the effect sizes from the two media sources.

3. Results

To examine overall single-observer consensus and accuracy aggregated across raters, we employed a random-effects meta-analysis. All analyses were performed using the R statistical package metafor (Viechtbauer, 2010). We estimated the amount of heterogeneity ($\tau^2$) and its standard error using a restricted maximum-likelihood estimator (REML); Cochran's Q-test was used to assess the significance of heterogeneity (Hedges & Olkin, 1985). We assessed publication bias using a regression model for the aggregate score that is typically seen for reliability measures in pri-

### Table 1

<table>
<thead>
<tr>
<th>Dimension</th>
<th>$k$</th>
<th>$M_z$ (SE)</th>
<th>95% CI</th>
<th>$t$</th>
<th>$r^2$ (SE)</th>
<th>$Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td>28</td>
<td>.32 (.04)</td>
<td>[.24, .39]</td>
<td>−0.10</td>
<td>0.01 (0.1)</td>
<td>18.29</td>
</tr>
<tr>
<td>Openness</td>
<td>26</td>
<td>.20 (.04)</td>
<td>[.12, .27]</td>
<td>0.73</td>
<td>0.01 (0.1)</td>
<td>16.80</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>26</td>
<td>.23 (.05)</td>
<td>[.15, .34]</td>
<td>−0.24</td>
<td>0.02 (0.02)</td>
<td>36.59</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>26</td>
<td>.26 (.05)</td>
<td>[.17, .35]</td>
<td>−0.24</td>
<td>0.01 (0.01)</td>
<td>27.55</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>27</td>
<td>.15 (.04)</td>
<td>[.07, .22]</td>
<td>0.46</td>
<td>0.01 (0.01)</td>
<td>16.06</td>
</tr>
<tr>
<td>Written communication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td>9</td>
<td>.30 (.05)</td>
<td>[.21, .40]</td>
<td>−0.27</td>
<td>0.00 (0.01)</td>
<td>2.91</td>
</tr>
<tr>
<td>Openness</td>
<td>8</td>
<td>.21 (.05)</td>
<td>[.11, .31]</td>
<td>0.25</td>
<td>0.00 (0.01)</td>
<td>3.63</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>8</td>
<td>.26 (.07)</td>
<td>[.13, .39]</td>
<td>0.09</td>
<td>0.01 (0.01)</td>
<td>4.88</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>8</td>
<td>.26 (.05)</td>
<td>[.16, .35]</td>
<td>−0.27</td>
<td>0.00 (0.01)</td>
<td>1.91</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>9</td>
<td>.21 (.05)</td>
<td>[.12, .30]</td>
<td>0.59</td>
<td>0.00 (0.01)</td>
<td>3.44</td>
</tr>
<tr>
<td>OSN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extraversion</td>
<td>19</td>
<td>.33 (.05)</td>
<td>[.22, .43]</td>
<td>−0.01</td>
<td>0.01 (0.01)</td>
<td>15.05</td>
</tr>
<tr>
<td>Openness</td>
<td>18</td>
<td>.19 (.05)</td>
<td>[.09, .30]</td>
<td>0.70</td>
<td>0.01 (0.01)</td>
<td>13.00</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>18</td>
<td>.24 (.07)</td>
<td>[.10, .37]</td>
<td>−0.27</td>
<td>0.03 (0.02)</td>
<td>31.67</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>18</td>
<td>.27 (.06)</td>
<td>[.14, .39]</td>
<td>−0.19</td>
<td>0.02 (0.02)</td>
<td>25.40</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>18</td>
<td>.11 (.05)</td>
<td>[.01, .20]</td>
<td>0.30</td>
<td>0.01 (0.01)</td>
<td>10.19</td>
</tr>
</tbody>
</table>

1. $p < .10.$
2. $p < .01.$

3.1. Consensus

All of the effect sizes for consensus were positive values, indicating significant agreement among the judges both within and across samples (Table 1). Although these correlations might appear small compared to common indices of inter-rater reliability, it is important to note that they represent the correlation between a single observer and all of the other raters, rather than the overall aggregate score that is typically seen for reliability measures in primary sources of empirical research (e.g., Cronbach’s $z$). Furthermore, although the aggregate effects in the current sample were all positive, the lower bounds of the 95% confidence intervals actually demarcated a small effect size in all cases except neuroticism. Thus, we suggest that conservative researchers might want to use the lower bound of the confidence interval when estimating the number of raters (see also Funder et al., 2013). Assessments of publication bias showed little evidence of bias in the current sample. All of the consensus effect sizes were relatively homogeneous and consistent across studies, indicating an absence of inter-study moderator variables and suggesting that studies’ authors were likely concerned with achieving a targeted level of inter-rater reliability when they were conducting the research.

3.2. Accuracy

The mean reported effect sizes for all five traits were positive (Table 2). As with the effects for consensus, the confidence intervals were quite large. The confidence intervals for all traits except neuroticism were greater than zero, however. Notably, the lower limit of the 95% confidence interval for extraversion was considerably higher than the other traits, which suggests greater confidence about the legibility of extraversion from written media and OSN profiles. Although the confidence interval for neuroticism included zero and was negative at its lower bound, most of the interval’s range was positive. The accuracy effects for neuroticism therefore appear to be less reliable compared to the other four Big Five traits.

Examination of the funnel plot asymmetry revealed little evidence for publication bias. However, the results of the heterogeneity analysis suggested a considerable mix in the magnitude of effect sizes for all traits except conscientiousness. Thus, it is likely that cross-study moderators might be able to explain the variability among the effects reported in the literature. Although we were restricted in our ability to examine such moderators due to the relatively small sample of effects available, the contrast between effects based on written text ($n = 9$) versus OSN profile information ($n = 21$) showed little evidence of differences for any of the individual traits, all $Zs < 1.13, ps > .11.$

4. Discussion

Not only do people tend to agree with each other when judging others’ personality traits from their written text and OSN, but these judgments might also be accurate for some traits. Here, we found evidence for the legibility of several Big Five traits from individuals’ self-generated written text and OSN profiles. Our findings were in line with models of incidental personality judgment (Gosling et al., 2002) and demonstrate that communication environments could be one medium considered within the framework of behavioral residue and identity claims. As such, when people write or create their online profile, they are making various identity claims. The content itself (e.g., grammar and style, posts), on the other hand, could be interpreted as behavioral residue. Most important, however, it seems that the general results across these studies suggest accurate effects for perceiving extraversion, openness to experience, agreeableness, and conscientiousness. Although the effects...
for neuroticism were generally positive, the data are less convincing that neuroticism can be judged accurately from OSN or written communications.

Across the studies and personality traits examined, we found that consensus among judges was generally positive. This suggests good reliability in the measurements made by the studies' authors. Although we did not find any evidence for variability in judges' consensus, researchers might want to examine some moderators of consensus in OSN and writing previously reported in the literature on interjudge perceptions. For example, John and Robins (1993) discussed how the evaluativeness of traits, the observability of traits, and social desirability could have an effect on interjudge agreement. Because both writing and OSN present a communication context, these variables might be especially relevant in such domains; researchers might therefore benefit from examining these variables in future work.

Furthermore, we found that accuracy estimates varied among the Big Five traits. Although we tested the source of information (writing versus OSN) as a potential moderator of accuracy, we did not find significant differences in the effect sizes as a function of medium. Still, a formal comparison across channels of media (e.g., photos, videos, writing) may be relevant for better understanding the relative contributions of various sources to accurate personality judgment. Of particular interest could be instances in which photos and writing appear to express conflicting information, raising the question of which source judges would favor in drawing their conclusions. Another moderator that could warrant further exploration is the evaluation of writing and OSN by close acquaintances versus complete strangers (Kenny, 1991; Paulhus & Reynolds, 1995). In this case, researchers might be able to estimate the degree to which perceptions of personality from written-materials and OSN are biased by interpersonal acquaintance.

Similarly, judges' sex, age, race, and cultural environment could have an effect on perceptions of personality in OSN. For example, research has suggested that people in East Asian cultures tend to think more holistically about the surrounding world and focus more attention on background information in photos compared to people in Western societies (e.g., Masuda & Nisbett, 2001). Should this extend to OSN, it is likely that different patterns of personality judgment may be observed between East Asian and North American raters. Similarly, future work may wish to investigate the role that racial group membership plays on the legibility and judgment of personality (Cosling et al., 2002). This seems like a potentially fertile area for future research, as written samples may be blinded to the perceptual cues that typically render the perception of race to be rather obvious. Comparisons between assessments of personality from cues in self-expressive and communicative media versus appearance may therefore be particularly informative for understanding the interaction between race and inferences of personality. Examination of these variables could be valuable for work on race perception, prejudice, and stereotypes.

Related to the effects of demographic variables, sampling context might also impact judgments of personality. Judgments might be enhanced for stimuli written or created in a lab setting versus those culled from more naturally-occurring sources in the outside world. Such an effect might suggest important differences in the behavior that manifests in more controlled settings versus those in contexts that might have higher ecological validity. Similarly, and not surprisingly, the more that targets write, the more accurately they may be judged in terms of extraversion, as sociability behavior that manifests in more controlled settings versus those culled from more naturally-occurring sources in the outside world (e.g., Kanagawa, Cross, & Markus, 2001). Additional exploration of these factors could be illuminative and raise important considerations for future researchers embarking on new work in this area.

5. Conclusion

In sum, the current work quantitatively reviewed the literature examining perceptions of personality from OSN and writing samples. We found that perceivers generally agree with each other when forming impressions of the Big Five traits from these sources and we hope that the estimates from the current analyses will be useful to researchers wishing to achieve high inter-rater reliabilities for judgments of personality from such media. Importantly, the data inspire high confidence that extraversion, conscientiousness, agreeableness, and openness to experience are judged...
accurately from self-expression materials. However, the data are less inspiring about the validity and reliability for judgments of personality. With the increased presence of individuals through their presence in online and written communication, both within and beyond perceptions of personality, may be productive for understanding both practical and conceptual aspects of social interactions and life outcomes.

References

References marked with an asterisk indicate articles included in the meta-analysis


Studies included in the analysis and not mentioned in the text


• van de Ven, N., Verschuren, A., & Denissen, J. (2013). Profiles at job-related social networking sites can be used to predict personality (Unpublished manuscript).

Weidman, A. C. & Levinson, C. A. (2013). I'm socially anxious, can you tell? Offline features of social anxiety manifest and are perceived in Facebook profiles. (Unpublished manuscript).

Weidman, A. C., Cheng, J. T., Chisolm, C., & Tracy, J. L. (2013). How can you tell if he's the one? Consensus, cue utility, and accuracy in personality judgments from online personal advertisements. (Unpublished manuscript).