

Corrigendum: Birds of a Feather Do Flock Together: Behavior-Based Personality-Assessment Method Reveals Personality Similarity Among Couples and Friends

Original article: Youyou, W., Schwartz, H. A., Stillwell, D., & Kosinski, M. (2017). Birds of a feather do flock together: Behavior-based personality-assessment method reveals personality similarity among couples and friends. *Psychological Science*, *28*, 276–284. doi:10.1177/0956797616678187

As a result of an oversight, the order in which the authors of this article were listed was incorrect. The correct order of authorship is as follows:

Wu Youyou, David Stillwell, H. Andrew Schwartz, and Michal Kosinski

Birds of a Feather Do Flock Together: Behavior-Based Personality-Assessment Method Reveals Personality Similarity Among Couples and Friends

Wu Youyou^{1,2}, David Stillwell³, H. Andrew Schwartz⁴, and Michal Kosinski⁵

¹Department of Psychology, University of Cambridge; ²Kellogg School of Management, Northwestern University; ³Judge Business School, University of Cambridge; ⁴Department of Computer Science, Stony Brook University; and ⁵Graduate School of Business, Stanford University

Psychological Science
2017, Vol. 28(3) 276–284
© The Author(s) 2017
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797616678187
www.psychologicalscience.org/PS
SAGE

Abstract

Friends and spouses tend to be similar in a broad range of characteristics, such as age, educational level, race, religion, attitudes, and general intelligence. Surprisingly, little evidence has been found for similarity in personality—one of the most fundamental psychological constructs. We argue that the lack of evidence for personality similarity stems from the tendency of individuals to make personality judgments relative to a salient comparison group, rather than in absolute terms (i.e., the *reference-group effect*), when responding to the self-report and peer-report questionnaires commonly used in personality research. We employed two behavior-based personality measures to circumvent the reference-group effect. The results based on large samples provide evidence for personality similarity between romantic partners ($n = 1,101$; $r_s = .20-.47$) and between friends ($n = 46,483$; $r_s = .12-.31$). We discuss the practical and methodological implications of the findings.

Keywords

similarity, personality assessment, reference-group effect, social network, close relationships

Received 5/4/15; Revision accepted 10/16/16

It is well established that in close relationships, individuals tend to be similar in a wide range of characteristics (McPherson, Smith-Lovin, & Cook, 2001), including age, education, race, religion, attitudes, and general intelligence (Rushton & Bons, 2005). Surprisingly, little evidence has been found for personality—a fundamental psychological construct that underpins much of the variation in human behaviors. Most past research has shown no or only weak similarity in personality between partners and between friends (Altmann, Sierau, & Roth, 2013; Anderson, Keltner, & John, 2003; Beer, Watson, & McDade-Montez, 2013; Botwin, Buss, & Shackelford, 1997; Buss, 1984a; Funder, Kolar, & Blackman, 1995; Rushton & Bons, 2005; Watson, Beer, & McDade-Montez, 2014; Watson, Hubbard, & Wiese, 2000a; Watson et al., 2004), with occasional findings indicating moderate similarity in the Big Five factors of Openness to

Experience and Conscientiousness between romantic partners (Donnellan, Conger, & Bryant, 2004; McCrae et al., 2008; Watson, Hubbard, & Wiese, 2000b). This has led researchers to maintain the conclusion drawn by an early theorist that “mating is essentially random for personality differences” (Eysenck, 1990, p. 252).

We argue that the lack of consistent evidence for personality similarity among couples or friends stems from the reliance on self-report and peer report¹ of personality in a majority of previous studies. These assessment methods are unsuitable for studying the similarity effect, because they are affected by a tendency of the respondents to

Corresponding Author:

Wu Youyou, Northwestern Institute on Complex Systems, Chambers Hall, 600 Foster St., Evanston, IL 60208-4057
E-mail: yw341@cam.ac.uk

judge themselves relative to a salient comparison group, rather than in absolute terms (the *reference-group effect*; Heine, Buchtel, & Norenzayan, 2008; Heine, Lehman, Peng, & Greenholtz, 2002). For instance, an introverted engineer might perceive himself as relatively extraverted if he is surrounded by a group of even more introverted engineer friends. The same bias affects peer report as well; the introverted friends of the engineer might also see him as extraverted by comparing him with themselves. In fact, some widely used personality questionnaires specifically instruct people to describe themselves “in relation to other people you know” (e.g., the International Personality Item Pool, IPIP, measuring the five-factor model of personality; Goldberg et al., 2006).

Several studies have found that self-reports of personality do not always correspond with behavioral measures (Heine et al., 2008; Ramírez-Esparza, Mehl, Álvarez-Bermúdez, & Pennebaker, 2009). The authors of these studies have suggested that the reference-group effect is a possible explanation. Subsequent experimental studies confirmed that the reference-group effect indeed pertains to questionnaire-based personality judgments (Credé, Bashshur, & Niehorster, 2010; Wood, Brown, Maltby, & Watkinson, 2012). We therefore argue that self- and peer report are inappropriate methods for studying personality similarity, because they amplify the differences in actual personality and obscure the similarity among partners and friends, who likely unconsciously treat one another as reference groups.

Indeed, rare evidence of personality similarity emerged from a few studies relying on personality measures that are less susceptible to the reference-group effect. For example, Botwin et al. (1997) and Buss (1984a) measured personality using independent interviewers' ratings and found similarity among spouses. Admittedly, this type of measure is still subject to the reference-group effect because the interviewer has his or her own reference group, but it affects both dyad members equally and therefore does not obscure the similarity between them. Buss (1984b) also found similarity between romantic partners by measuring personality using self- and peer-reported frequencies of certain personality-related behaviors (the *act-frequency approach*; Buss & Craik, 1983). Introversion, for example, was assessed by asking participants to judge whether in the last 3 months they “watched the soap opera on TV” or “went for a long walk alone” (Buss, 1984b, p. 368). This approach focuses on concrete behaviors and thus leaves less room for subjective comparisons.

In light of these mixed findings, we aimed to address the reference-group effect and reexamine the existence of personality similarity between romantic couples and

between friends. We employed two behavior-based personality measures to circumvent the reference-group effect.

The first approach measured personality using a common type of digital footprint: Facebook Likes. Facebook users generate Likes by clicking a Like button on Facebook Pages related to products, famous people, books, etc.² This feature allows users to express their preferences for a variety of content. It has been shown that Likes can be used to accurately assess people's personality (Kosinski, Stillwell, & Graepel, 2013; Youyou, Kosinski, & Stillwell, 2015). For example, people who score high on Extraversion tend to Like “partying,” “dancing,” and celebrities such as “Snooki” (a reality-TV personality).³

The second approach measured personality using digital records of language use: Facebook status updates. Facebook users write status updates to share their thoughts, feelings, and life events with friends. Previous research has consistently found links between personality and language use (Hirsh & Peterson, 2009; Mehl, Gosling, & Pennebaker, 2006; Tausczik & Pennebaker, 2010). Extraverts, for example, tend to use more words describing positive emotions (e.g., “great,” “happy,” or “amazing”; H. A. Schwartz et al., 2013) than introverts do. Several studies have demonstrated accurate personality assessment based on people's language use in social media (Farnadi et al., 2014; Sumner, Byers, Boochever, & Park, 2012), including Facebook status updates (Park et al., 2014; H. A. Schwartz et al., 2013).

For both Likes-based and language-based approaches, we measured personality in the following way. First, we obtained a sample of participants with both self-reports of personality and Facebook data. Next, we built a series of predictive models to link self-reports of personality with Likes or language use, respectively. This process allowed us to establish which digital signals were indicative of specific personality traits. The resulting models were then applied to a separate sample of romantic partners and friends to generate personality scores for these participants. The personality scores were correlated between dyad members to measure similarity.

Notably, although both Likes-based and language-based models are developed based on participants' self-reported personality scores, they do not inherit the reference-group effect. The reference-group bias that contaminates personality similarity is a result of individuals using different standards, norms, or reference groups to evaluate themselves. In our analysis, the same personality-prediction models were applied to the entire sample, and therefore the evaluation standards were uniform across all participants.

Method

Participants

This study relied on three samples obtained from the myPersonality project (<http://mypersonality.org>). MyPersonality was a popular Facebook application that allowed users to take psychological tests and receive feedback on their scores. A portion of participants provided opt-in consent to allow us to record their test scores and contents of their Facebook profile. The average participant age was 24.1 years. Females constituted 61.1% of the sample, and males constituted 38.9%.

Sample 1 was used to build the Likes-based personality-assessment models. It contained 295,320 participants who completed personality questionnaires and had at least 20 Likes on their Facebook profile. Sample 2 was used to develop the language-based personality-assessment models. It contained 59,547 participants who completed personality questionnaires and wrote at least 500 words across all of their status updates. Sample 3 was used to study the existence of personality similarity between romantic partners and between friends. It contained 247,773 individuals forming a total of 5,042 heterosexual romantic couples and 138,553 friendship dyads. Romantic couples were identified using the “relationship status” field of the Facebook profile, and friendship connections were identified using Facebook friend lists. To ensure that all dyads included in the analysis were independent from one another, we randomly chose one dyad for the individuals belonging to multiple friendship dyads.

Self-report of personality

Self-reports of personality were obtained using a 20- to 100-item IPIP questionnaire (Goldberg et al., 2006) measuring the widely accepted five-factor model of personality (Revised NEO Personality Inventory; Costa & McCrae, 1992). Reliability scores for the 100-item questionnaire, completed by 29.4% of the participants, were as follows—Openness to Experience: Cronbach’s $\alpha = .84$, Conscientiousness: $\alpha = .92$, Extraversion: $\alpha = .93$, Agreeableness: $\alpha = .88$, and Neuroticism: $\alpha = .93$. Corresponding values for the 20-item version, completed by 56.4% of the participants, were as follows—Openness to Experience: Cronbach’s $\alpha = .48$, Conscientiousness: $\alpha = .67$, Extraversion: $\alpha = .73$, Agreeableness: $\alpha = .58$, and Neuroticism: $\alpha = .65$. The remaining participants (14.2%) completed the IPIP questionnaire ranging from 30 to 90 items (in intervals of 10). Self-reports were available for all participants in Samples 1 and 2, and for 4,287 romantic couples and 103,329 friendship dyads in Sample 3.

Likes-based personality assessment

The Likes-based personality-assessment model was built using Sample 1 following the procedure described in detail in Youyou et al. (2015). We first transformed participants’ Like data into a matrix, in which each row represented a participant, and each column represented a Like. The (i, j) entry was set to 1 if participant i liked object j , and 0 otherwise. A substantial number of Likes were associated with only a few participants in this sample, and some participants had only a small number of Likes. Since the assessment models leveraged the association between liking certain things and having a particular personality type, it was necessary to have enough combinations of personality profiles and Likes as training examples. The matrix was therefore trimmed so that participants with fewer than 20 Likes and Likes associated with fewer than 20 participants were removed. The resulting matrix consisted of 295,320 participants (rows) and 148,128 unique Likes (columns).

For each of the five personality traits, a linear regression model was fitted to predict the self-reported personality scores from the participant-Like matrix (each column was treated as a variable); a combination of L1 (least absolute shrinkage and selection operator, or LASSO; Tibshirani, 1996) and L2 (ridge; Hoerl & Kennard, 1970) penalties were used for the models.⁴ A 10-fold cross-validation was applied in each model to avoid overfitting.

The resulting five models, one for each personality trait, were applied to a separate sample of romantic couples ($n = 990$) and friends ($n = 41,880$) in Sample 3, to generate behavior-based personality scores for these participants. This sample contained only couples and friendship dyads in which both members had at least 20 Likes on their profile. The average number of Likes per participant was 159.4. We removed Likes shared between each pair of dyad members to ensure that they did not artificially inflate similarity. Such an overlap in Likes between dyad members was relatively low: friends, on average, shared 5.2 Likes, or 1.4% of their joint Likes; romantic partners shared 12.8 Likes, or 3.5% of their joint Likes.

To evaluate the predictive accuracy of the Likes-based models, we correlated Likes-based and self-reported personality scores for a subset of participants in Sample 3 (note that the model was developed using Sample 1). After shared Likes were removed for all individuals, the correlations were as follows—Openness to Experience: $r(22,692) = .39$, Conscientiousness: $r = .28$, Extraversion: $r = .31$, Agreeableness: $r = .25$, and Neuroticism: $r = .29$.

Language-based personality

The language-based personality-assessment model was developed using Sample 2, with an open-vocabulary

approach similar to the one employed by Park et al. (2014). We first extracted words and phrases (i.e., sequence of words) from participants' status updates and then transformed them into two types of predictors: (a) binary indicators of whether the participant used each word and (b) relative frequencies of each word or phrase (as compared with the total number of words that each participant wrote). Words and phrases used by less than 1% of the participants were excluded when we created the predictors. The two types of predictors were each represented as a matrix and underwent randomized principal-component analysis (RPCA; Martinsson, Rokhlin, & Tygert, 2011) independently. They were then combined into a single participant-language matrix. For each of the five personality traits, a linear regression model with an L2 (ridge) penalty was fitted to predict the self-reported personality scores from the participant-language matrix. A 10-fold cross-validation was applied in each model to avoid overfitting.

Because words and phrases shared between two partners and between friends could artificially inflate personality similarity, it was necessary to control for the overlap in language between dyad members. However, we could not exclude all the words shared between dyad members (as we did with Likes), because most of the common words would be removed as a result, which would lower the predictive accuracy. Instead, we randomly split all the available words and phrases into two halves, and submitted each half to the procedures described in the previous paragraph. The two resulting matrices were regressed onto participants' self-reported personality scores to build two independent sets of predictive models. Finally, the two sets of models were each applied to a different member of the dyad. This process ensured that even if two dyad members used the same words or phrases, the two different models applied to each of them separately would capture only distinct parts of the overlap.

The two sets of models developed here were applied to 282 romantic couples and 5,674 friendship dyads in Sample 3. These dyads all consisted of members that both had status updates available and wrote at least 500 words across all of their status updates. Participants in this sample wrote 4,474 words on average.

To determine the predictive accuracy of the language-based models, we correlated language-based and self-reported personality scores for a subset of participants in Sample 3 (note that the model was developed using Sample 2). After an overlap in language was controlled for (by applying independent models to each dyad member), the correlations were as follows—Openness to Experience: $r(2,718) = .37$, Conscientiousness: $r = .32$, Extraversion: $r = .34$, Agreeableness: $r = .30$, and Neuroticism: $r = .33$.

Measuring similarity

Similarity between dyad members was measured by correlating their scores on a given personality trait across all dyads. Correlations were calculated for self-report, Likes-based, and language-based measures, respectively, between partners and between friends. Additionally, we calculated correlations between one dyad member's Likes-based score and this person's partner's or friend's language-based score across all dyads.

For romantic couples, personality scores were aligned by gender, and Pearson product-moment correlation coefficients were used. For friendship dyads, intraclass correlations were used because dyads cannot be aligned by gender, and the assignment of dyad members as person A or person B is arbitrary (see Watson et al., 2000b).⁵

Results

The goal of this study was to examine the degree of personality similarity between romantic partners and between friends. The results based on three different personality measures are presented in Figure 1.

The Likes-based scores between dyad members showed significant positive correlations across all five personality traits—romantic couples: mean $r(988) = .24$, 95% confidence interval (CI) = [.18, .30]; friends: mean $r(83,758) = .14$, 95% CI = [.13, .15].⁶ An even stronger effect was observed in the language-based results—romantic couples: mean $r(280) = .38$, CI = [.28, .48]; friends: mean $r(11,346) = .24$, 95% CI = [.22, .26]. Using both Likes-based and language-based measures, the correlations did not differ substantially between same-sex and opposite-sex friendships (all differences were .03 or less). In contrast, self-reports showed weak to negligible personality similarity for both romantic couples, mean $r(4,285) = .10$, 95% CI = [.07, .13], and friends, mean $r(206,656) = .06$, 95% CI = [.06, .07]. All these correlations were significant at $p < .001$.

The strength of personality similarity became clear when compared with the similarity observed for other variables. Personality similarity was not as strong as similarity in age—romantic couples: $r(2,458) = .81$, 95% CI = [.80, .82], $p < .001$; friends: $r(85,076) = .57$, 95% CI = [.57, .58], $p < .001$. However, it was comparable to or stronger than similarity in IQ: $r(550) = .21$, 95% CI = [.13, .29], $p < .001$ (this sample included both romantic couples and friends because there were not enough romantic couples in which both partners had IQ scores, $n = 44$, to allow for a meaningful comparison).

These similarity results were based on personality scores measured using nonoverlapping Likes and language features. This was because the overlap in Likes or

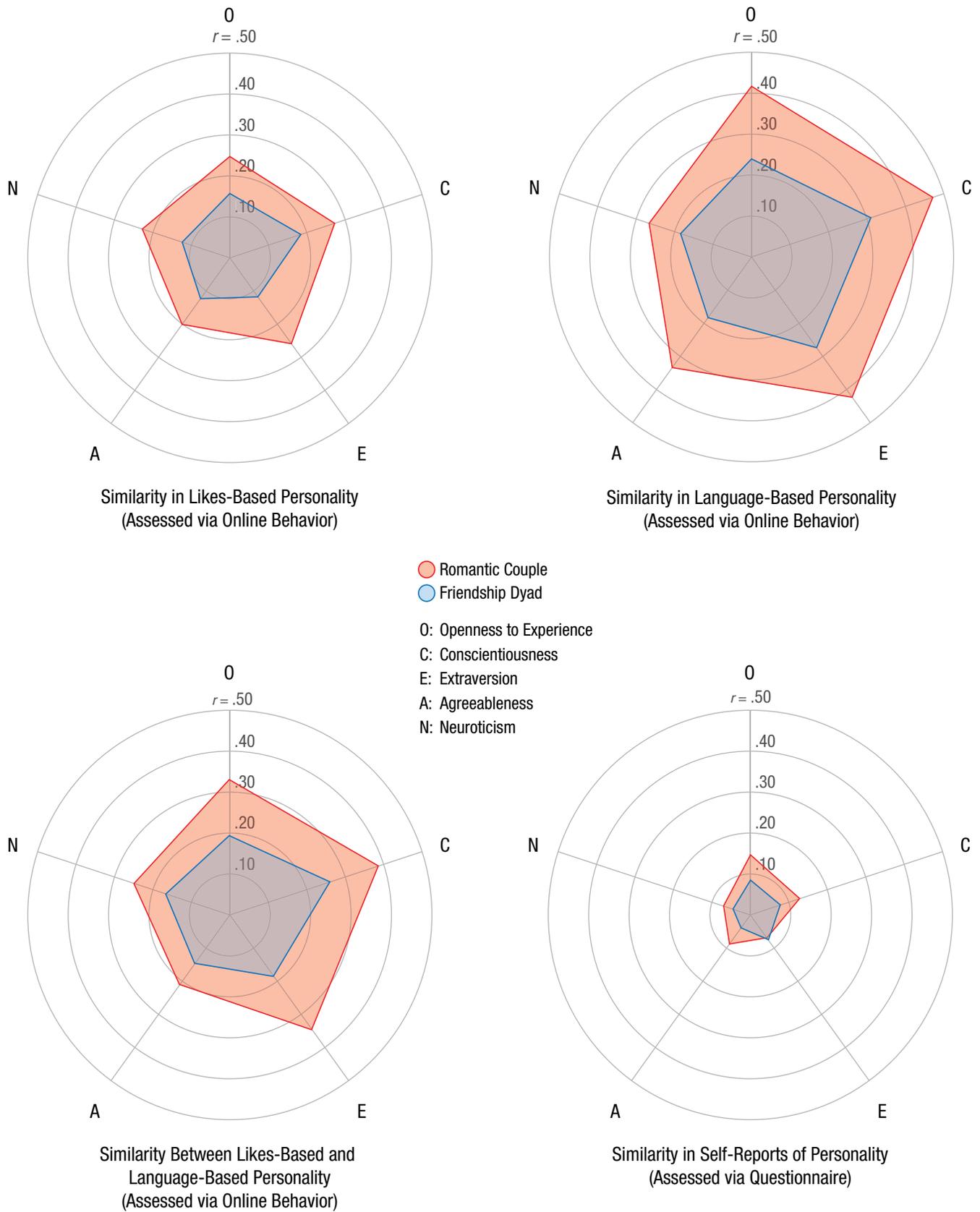


Fig. 1. Radar charts showing the similarity in the Big Five personality traits between romantic partners and between friends. Results are shown separately for analyses based on Facebook Likes, Facebook language use, the combination of these two measures, and self-report questionnaires.

language features between dyad members might have been driven by factors other than personality, such as a shared environment, shared culture, or interpersonal influence. However, it also might have been partially driven by actual personality similarity. Consequently, these results represent a lower-bound estimate of similarity—some effect was lost. To calculate the upper-bound estimate, we performed the same analyses without controlling for shared Likes or language features. As expected, the results showed a stronger level of similarity. The Likes-based correlations were as follows—romantic couples: mean $r(1,082) = .33$, 95% CI = [.28, .38]; friends: mean $r(87,842) = .19$, 95% CI = [.18, .20]; the language-based ones were as follows—romantic couples: mean $r(280) = .41$, 95% CI = [.31, .50]; friends: mean $r(11,346) = .25$, 95% CI = [.23, .27], all $ps < .001$.

One potential problem with the preceding analyses was that the scores of both dyad members were based on the same type of data, namely, Likes or status updates. This was problematic, as Facebook's News Feed and its recommendation system might cause an artificial covariation between friends' Likes or status updates (i.e., *common-method bias*; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). For example, Facebook recommends Pages to its users that are similar to the ones their friends liked. Also, users are constantly exposed to their friends' status updates in the News Feed and are therefore prone to post about similar topics. While we already controlled for overlap in Likes and language features, to further reduce potential sources of bias, we correlated the Likes-based scores of one person with the language-based scores of that person's partner or friend. The results were similar to the ones already reported: The average personality similarity across the five traits was as follows—romantic couples: mean $r(1,055) = .31$, 95% CI = [.26, .36], $ps < .001$; friends: mean $r(32,552) = .19$, 95% CI = [.18, .20], $ps < .001$.

Additionally, a series of analyses was performed to rule out alternative explanations for the observed personality similarity. First, we calculated the correlations between random pairs of participants to gauge the baseline similarity between strangers. None of the correlations were significant ($|r_s| < .01$) for random dyads.

Second, dyad members' scores were correlated controlling for the number of Likes that they had or the number of words written in the status updates. This was to ensure that similarity in predicted personality between partners or between friends was not due to having a similar number of digital signals. These partial correlations were very similar to the zero-order ones (within .02 of the original values).

Third, we investigated the extent to which the observed personality similarity was a by-product of similarity in other traits (Buss, 1984a). To this end, we reran the

analyses while controlling for dyad members' age and education. For education, we calculated correlations for a subsample of friendship dyads in which both members were college graduates. We used keywords such as "university" or "college" (excluding "community college") in the school names shown on Facebook profiles to identify a sample of participants with higher education degrees. The level of similarity between friends did not change considerably after taking education into account: All correlations were within .04 of the original values. Unfortunately, information about education was not available for enough of the romantic couples to provide for a meaningful comparison ($n = 17$ for the Likes-based approach and $n = 6$ for the language-based approach). The analysis was therefore limited to friends only.

Similarly, little change was observed when we controlled for age. For both romantic couples and friends, partial correlations that controlled for age were all within .03, compared with zero-order ones. The only exception was Conscientiousness, for which the correlations decreased on average by .10 for romantic couples and .08 for friends across the three methods. Nevertheless, the similarity in Conscientiousness remained significant at $p < .001$ (romantic partners: mean $r = .30$; friends: mean $r = .16$).

Discussion

Our findings provide evidence that romantic partners as well as friends are characterized by similar personalities. We measured personality traits relying on three different sources of data: traditional self-report questionnaires, digital records of behaviors and preferences, and language use. Relatively strong similarity was detected between romantic partners and between friends when we used Likes-based and language-based measures. By contrast, self-reports yielded only weak to negligible similarity. Across all three methods, stronger personality similarity was found for romantic couples than for friends.

We also showed that dyadic similarity in most personality traits was unlikely to be driven simply by similarity in age or education. The only exception was dyad members' similarity in Conscientiousness, which was partially explained by their similarity in age. Compared with the other four traits, Conscientiousness is most strongly positively associated with age, especially before the age of 30 (Donnellan & Lucas, 2008; Soto, John, Gosling, & Potter, 2011). Because 88% of the participants were between 18 and 30 years old, it is not surprising that partners' and friends' similarity in Conscientiousness was partially due to their similarity in age.

In which of the five personality traits were romantic partners and friends most similar? After controlling for age, we found that Openness to Experience displayed

the strongest similarity in self-reports, Likes-based results, and Likes-language correlations for both romantic couples and friendship dyads. Language-based results, however, showed the strongest effect in Extraversion. However, we cannot draw definitive conclusions on the basis of our present analysis, because (a) the patterns were not consistent across all the methods that we employed, and (b) the effect sizes could be influenced by several factors, such as the strength of the reference-group effect, the accuracy of the assessment models, and common method bias. These factors might affect the five traits differently and to varying degrees.

Together, these results challenge the widely accepted notion that individuals in close relationships are not similar in personality. We argue that the scarcity of the evidence for the similarity effect is likely due to the reference-group effect. Notably, our results are consistent with those obtained in rare previous studies that relied on personality-assessment methods resistant to the reference-group effect (Botwin et al., 1997; Buss, 1984a, 1984b). On the other hand, the fact that the results presented contradict the majority of previous findings means that they should be treated with caution. We hope that future research will replicate our findings using other methods.

From a methodological perspective, the present research highlights the limitations of questionnaire measures (Heine et al., 2002). While self-report personality questionnaires provide excellent reliability and validity in most applications (Costa & McCrae, 1992; Goldberg et al., 2006; Ozer & Benet-Martínez, 2006), they fail to assess personality similarity between individuals. As illustrated in the present research, personality assessment based on digital records of preferences and language, while still relatively new and unproven, has the potential to address this issue.

Both Likes-based and language-based personality-assessment methods are unlikely to be affected by the reference-group effect: People do not use words or do things just because their friends refrained from doing so. In fact, the reverse effect is likely: A shared environment, culture, or interpersonal influence may inflate the similarity in Likes and language between dyad members. For example, people from the same cohort are likely to be fans of similar pop stars of their generation, and two friends might both write “explosion” in their status updates because they live in the same area and an explosion recently happened nearby. We addressed these limitations by controlling for age and education, removing Likes shared between dyad members, applying two disjoint sets of language-based models to each of the dyad members, and correlating Likes-based scores with language-based scores. However, we might have still omitted variables,

such as subcultures, driving the adoption of similar language patterns or preferences online.

Finally, if the reference-group effect obscures the similarity in personality measured using self-reports, why has similarity been consistently detected for self-reported attitudes and values, such as religious and political views (e.g., Alford, Hatemi, Hibbing, Martin, & Eaves, 2011; Gaunt, 2006)? There are two possible explanations. First, the similarity in values and attitudes might have been underestimated and would be higher if the reference-group effect was eliminated. Past research shows that larger effects were generally found when behavioral or objective indicators of attitudes were used (e.g., church attendance; see Alford et al., 2011; Gaunt, 2006; Watson et al., 2004) compared with self-ratings on broad statements using the Likert scale (e.g., “How religious are you?”; see Caspi, Herbener, & Ozer, 1992; Gaunt, 2006; Lee et al., 2009; Watson et al., 2014). This pattern suggests that the reference-group effect applies to self-reported attitudes as well. Second, self-reports of values and attitudes might be subject to the counter-motive of conformity, canceling the reference-group effect. Past studies have found that people shift their attitudes to align with those of their romantic partners (Davis & Rusbult, 2001; Kalmijn, 2005). In contrast, no such tendency has been discovered for personality (Anderson et al., 2003; Caspi et al., 1992). We hope that future research will replicate our analysis in attitudinal domains, especially in basic values (e.g., S. H. Schwartz, 1992), a domain in which measures other than self-report are in short supply.

Action Editor

Ralph Adolphs served as action editor for this article.

Author Contributions

W. Youyou designed the research. M. Kosinski and D. Stillwell collected and cleaned the data, and W. Youyou, H. A. Schwartz, and M. Kosinski performed the analysis. W. Youyou, D. Stillwell, H. A. Schwartz, and M. Kosinski wrote the manuscript.

Acknowledgments

We thank John Rust, Patrick Morse, Sandra Matz, Jason Rentfrow, Josh Sacco, Vesselin Popov, and Jingwei Yu for their critical reading of the manuscript. We thank Isabelle Abraham for proofreading.

Declaration of Conflicting Interests

D. Stillwell received revenues as the owner of the myPersonality application that collected the data used in this research. The authors declared that they had no other potential conflicts of interest with respect to their authorship or the publication of this article.

Notes

1. Peer report here refers to when members in a dyad (e.g., a couple or a pair of friends) rate each other's personality.
2. Although Facebook users can Like many different types of content on Facebook—including status updates, photos, comments, pages, and links—we measured Likes associated only with Facebook Pages.
3. See Kosinski et al. (2013) for more examples of Likes strongly associated with each personality trait.
4. LASSO and ridge are variable-selection-and-constraint methods that are best used for a large number of predictors, especially when they are highly correlated. They constrain, or penalize, the absolute size of the regression coefficients so that no coefficients are too large when compared with others.
5. We also tested an alternative approach: randomly assigning friendship dyad members as person A or person B, and conducting Pearson product-moment correlations. The results were very similar to those obtained with the intraclass correlation (within .02 of the original values).
6. The average correlations across the five traits were calculated using Fisher's r -to- z transformation.

References

- Alford, J. R., Hatemi, P. K., Hibbing, J. R., Martin, N. G., & Eaves, L. J. (2011). The politics of mate choice. *The Journal of Politics, 73*, 362–379.
- Altmann, T., Sierau, S., & Roth, M. (2013). I guess you're just not my type. *Journal of Individual Differences, 34*, 105–117.
- Anderson, C., Keltner, D., & John, O. P. (2003). Emotional convergence between people over time. *Journal of Personality and Social Psychology, 84*, 1054–1068.
- Beer, A., Watson, D., & McDade-Montez, E. (2013). Self-other agreement and assumed similarity in neuroticism, extraversion, and trait affect: Distinguishing the effects of form and content. *Assessment, 20*, 723–737.
- Botwin, M. D., Buss, D. M., & Shackelford, T. K. (1997). Personality and mate preferences: Five factors in mate selection and marital satisfaction. *Journal of Personality, 65*, 107–136.
- Buss, D. M. (1984a). Marital assortment for personality dispositions: Assessment with three different data sources. *Behavior Genetics, 14*, 111–123.
- Buss, D. M. (1984b). Toward a psychology of person-environment (PE) correlation: The role of spouse selection. *Journal of Personality and Social Psychology, 47*, 361–377.
- Buss, D. M., & Craik, K. H. (1983). The act frequency approach to personality. *Psychological Review, 90*, 105–126.
- Caspi, A., Herbener, E. S. S., & Ozer, D. J. J. (1992). Shared experiences and the similarity of personalities: A longitudinal study of married couples. *Journal of Personality and Social Psychology, 62*, 281–291.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual* (Vol. 4). Odessa, FL: Psychological Assessment Resources.
- Credé, M., Bashshur, M., & Niehorster, S. (2010). Reference group effects in the measurement of personality and attitudes. *Journal of Personality Assessment, 92*, 390–399.
- Davis, J. L., & Rusbult, C. E. (2001). Attitude alignment in close relationships. *Journal of Personality and Social Psychology, 81*, 65–84.
- Donnellan, M. B., Conger, R. D., & Bryant, C. M. (2004). The Big Five and enduring marriages. *Journal of Research in Personality, 38*, 481–504.
- Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the Big Five across the life span: Evidence from two national samples. *Psychology and Aging, 23*, 558–566.
- Eysenck, H. J. (1990). Genetic and environmental contributions to individual differences: The three major dimensions of personality. *Journal of Personality, 58*, 245–261.
- Farnadi, G., Sushmita, S., Sitaraman, G., Ton, N., De Cock, M., & Davalos, S. (2014). A multivariate regression approach to personality impression recognition of vloggers. *Proceedings of the 2014 ACM Multi Media on Workshop on Computational Personality Recognition* (pp. 1–6). Retrieved from <http://dl.acm.org/citation.cfm?id=2659526&dl=ACM&coll=DL&CFID=863560964&CFTOKEN=93316611>
- Funder, D. C., Kolar, D. C., & Blackman, M. C. (1995). Agreement among judges of personality: Interpersonal relations, similarity, and acquaintanceship. *Journal of Personality and Social Psychology, 69*, 656–672.
- Gaunt, R. (2006). Couple similarity and marital satisfaction: Are similar spouses happier? *Journal of Personality, 74*, 1401–1420.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84–96.
- Heine, S. J., Buchtel, E. E., & Norenzayan, A. (2008). What do cross-national comparisons of personality traits tell us? The case of conscientiousness. *Psychological Science, 19*, 309–313.
- Heine, S. J., Lehman, D. R., Peng, K. P., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales?: The reference-group effect. *Journal of Personality and Social Psychology, 82*, 903–918.
- Hirsh, J. B., & Peterson, J. B. (2009). Personality and language use in self-narratives. *Journal of Research in Personality, 43*, 524–527.
- Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*, 55–67.
- Kalmijn, M. (2005). Attitude alignment in marriage and cohabitation: The case of sex-role attitudes. *Personal Relationships, 12*, 521–535.
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences, USA, 110*, 5802–5805.
- Lee, K., Ashton, M. C., Pozzebon, J. A., Visser, B. A., Bourdage, J. S., & Ogunfowora, B. (2009). Similarity and assumed similarity in personality reports of well-acquainted persons. *Journal of Personality and Social Psychology, 96*, 460–472.
- Martinson, P. G., Rokhlin, V., & Tygert, M. (2011). A randomized algorithm for the decomposition of matrices. *Applied and Computational Harmonic Analysis, 30*, 47–68.

- McCrae, R. R., Martin, T. A., Hřebícková, M., Urbánek, T., Boomsma, D. I., Willemsen, G., & Costa, P. T., Jr. (2008). Personality trait similarity between spouses in four cultures. *Journal of Personality, 76*, 1137–1164.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*, 415–444.
- Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology, 90*, 862–877.
- Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology, 57*, 401–421.
- Park, G. J., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., . . . Seligman, M. E. P. (2014). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology, 108*, 934–952.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*, 879–903.
- Ramírez-Esparza, N., Mehl, M. R., Álvarez-Bermúdez, J., & Pennebaker, J. W. (2009). Are Mexicans more or less sociable than Americans? Insights from a naturalistic observation study. *Journal of Research in Personality, 43*, 1–7.
- Rushton, J. P., & Bons, T. A. (2005). Mate choice and friendship in twins: Evidence for genetic similarity. *Psychological Science, 16*, 555–559.
- Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., . . . Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLoS ONE, 8*(9), Article e73791. doi:10.1371/journal.pone.0073791
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1–65). San Diego, CA: Academic Press.
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology, 100*, 330–348.
- Sumner, C., Byers, A., Boochever, R., & Park, G. J. (2012). Predicting Dark Triad personality traits from Twitter usage and a linguistic analysis of tweets. In D. Tao, M. A. Wani, T. Khoshgoftaar, X. Zhu, & N. Seliya (Eds.), *Proceedings: 2012 11th International Conference on Machine Learning and Applications, ICMLA 2012* (Vol. 2, pp. 386–393). Los Alamitos, CA: IEEE.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*, 24–54.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society B: Statistical Methodology, 58*, 267–288.
- Watson, D., Beer, A., & McDade-Montez, E. (2014). The role of active assortment in spousal similarity. *Journal of Personality, 82*, 116–129.
- Watson, D., Hubbard, B., & Wiese, D. (2000a). General traits of personality and affectivity as predictors of satisfaction in intimate relationships: Evidence from self- and partner-ratings. *Journal of Personality, 68*, 413–449.
- Watson, D., Hubbard, B., & Wiese, D. (2000b). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology, 78*, 546–558.
- Watson, D., Klohnen, E. C., Casillas, A., Simms, E. N., Haig, J., & Berry, D. S. (2004). Match makers and deal breakers: Analyses of assortative mating in newlywed couples. *Journal of Personality, 72*, 1029–1068.
- Wood, A. M., Brown, G. D. A., Maltby, J., & Watkinson, P. (2012). How are personality judgments made? A cognitive model of reference group effects, personality scale responses, and behavioral reactions. *Journal of Personality, 80*, 1275–1311.
- Youyou, W., Kosinski, M., & Stillwell, D. J. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences, USA, 112*, 112–116.