

# Self–other agreement in personality development in romantic couples



European Journal of Personality

0(0) 1–17

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0890207020969011

journals.sagepub.com/home/ejop



Madeline R Lenhausen<sup>1</sup>, Manon A van Scheppingen<sup>2</sup> and  
Wiebke Bleidorn<sup>1</sup> 

## Abstract

A large body of evidence indicates that personality traits show high rank-order stability and substantial mean-level changes across the lifespan. However, the majority of longitudinal research on personality development has relied on repeated assessments of self-reports, providing a narrow empirical base from which to draw conclusions and develop theory. Here, we (1) tested whether self- and informant-reports provided by couples show similar patterns of rank-order stability and mean-level change and (2) assessed self–other agreement in personality development. We charted the Big Five personality trajectories of 255 couples ( $N = 510$ ;  $M$  age = 27.01 years) who provided both self- and partner-reports at four assessments across 1.5 years. Results indicated similar rank-order stabilities in self- and partner-report data. Latent growth curve models indicated no significant differences between self- and partner-reported personality trajectories, with exceptions to extraversion and agreeableness. We further found strong cross-sectional agreement across all Big Five traits and assessment waves as well as moderate self–other agreement in personality change in emotional stability and agreeableness. These findings highlight the relevance of multi-method assessments in personality development, while providing information about personality stability and change. Discussion focuses on the theoretical implications and future directions for multi-method assessments in longitudinal personality research.

## Keywords

self–other agreement, personality change, informant-report, Big Five, latent growth curve model

Received 27 May 2020; Revised 15 September 2020; accepted 23 September 2020

A large body of research has shown that personality traits are both relatively stable and yet malleable enough to undergo marked changes across the lifespan (Bleidorn & Hopwood, 2019; Roberts & Mroczek, 2008). Although the general patterns of stability and change in personality traits are well established, there are limitations to this body of work that undermine the generalizability of these findings. Most importantly, the vast majority of studies have relied on self-report measures to assess the stability of and change in personality traits. Self-report measures are the most common method to assess individual differences and can provide reliable and valid information about personality traits (Nave et al., 2018; Tracy et al., 2009). However, they are not without limitations and should ultimately be complemented by other measures (Bornstein, 2009).

Informant-reports—from friends, family members, or spouses—can provide a valuable additional

perspective on personality stability and change and are useful for validating and complementing self-report assessments (Hofmann et al., 2009; Leising et al., 2010). Compared to the large body of self-report research, relatively few studies have examined adult personality trait stability and change using informant-report data and even less is known about the degree to which self- and informant-reports provide corresponding estimates of personality trajectories.

<sup>1</sup>Department of Psychology, University of California, Davis, USA

<sup>2</sup>Department of Developmental Psychology, Tilburg University, the Netherlands

### Corresponding author:

Madeline R Lenhausen, Department of Psychology, University of California, Davis, 1 Shields Avenue, Davis, CA 95616, USA.

Email: lenhausen14@gmail.com

As such, several open questions remain: Do longitudinal assessments of informant-reports replicate patterns of stability and change in self-reports? Specifically, do analyses of rank-order stability and mean-level change in self- and informant-report data yield similar results? And do reports from the self and informants agree in their personality ratings over time? In this study, we begin to address these questions using self- and partner-reports from couples with and without children who reported on their own and their partner's Big Five personality traits (John et al., 2008)—emotional stability (neuroticism reversed), extraversion, openness, agreeableness, and conscientiousness—at four assessment waves over a period of 1.5 years.

### *Self-report research on personality stability and change*

Most research on personality development has focused on two types of change—rank-order stability and mean-level change. More recently, researchers have become increasingly interested in a third type of change—individual differences in change (Bleidorn et al., 2019; Mroczek & Spiro, 2003; Schwaba & Bleidorn, 2018). Research on rank-order stability—the degree to which the relative ordering of individuals on a trait is maintained over time—indicated that personality traits are relatively but not perfectly stable across the lifespan (Bleidorn & Hopwood, 2019). Meta-analytic evidence suggests that the rank-order stability of traits is lowest in young adulthood and increases across the lifespan, especially through middle adulthood (Roberts & DelVecchio, 2000; Robins et al., 2001).

Mean-level change refers to how much a population changes in a given trait on average over a certain period of time. Self-report research has repeatedly found that most people increase in traits that reflect greater social maturity such as emotional stability, conscientiousness, and agreeableness, especially during the period of young adulthood (Bleidorn, 2012; Neyer & Lehnart, 2007; Roberts & Mroczek, 2008; Roberts et al., 2006, 2008; Robins et al., 2001; Soto et al., 2011; Specht et al., 2011, 2014). This trend seems to generalize across cohorts, genders, and, to a certain extent, also across samples from different cultures (Bleidorn et al., 2013; Soto et al., 2011; Wortman et al., 2012). Referring to young adults' increasing capacity to become a productive contributor to society, this pattern has been termed the *maturity principle* of personality development (Roberts & Mroczek, 2008; Roberts & Nickel, in press).

Individual differences in change reflect individual deviations from the normative trajectory, and are often expressed as variance around a mean-level trend in a population (e.g. Allemand et al., 2007). That is, people are not necessarily changing at the same rate in the same direction, but differ in their

individual trajectories. Recent studies found that individual differences in personality trait change are present throughout the lifespan (Schwaba & Bleidorn, 2018) but most pronounced during emerging and young adulthood (Pusch et al., 2019).

These findings appear to provide a solid foundation upon which scholars can build on to develop theories of personality stability and change. However, the reliance on self-report questionnaires has limited the conclusions we can draw from this literature. It is well established that any single source of personality assessment is afflicted with specific limitations and disadvantages (Bornstein, 2009; Kubinger & Litzenberger, 2006). Most important for the present study, both stability and change in personality traits can be influenced by rater-specific biases. In fact, longitudinal behavioral genetic research found that rater-specific method effects are both heritable and relatively stable suggesting that estimates of personality stability may be confounded with stable rater-specific effects (Kandler et al., 2010). For example, research indicated that people tend to use behavioral cues when judging others' personality but tend to rely on implicit knowledge when reporting on their own personality (Hofmann et al., 2009). As such, self-reports may over-emphasize personality stability because of how unwavering implicit aspects of personality are as compared to other methods that may be more tailored to assess changes in traits (Bleidorn et al., 2020).

To quantify and address potential validity threats to self-reports, researchers have used other sources of personality data. Much of this research has relied on informant-reports which can be used to validate self-report findings by assessing whether or not they replicate the normative trends found in self-reports (Vazire, 2006). However, relatively little *longitudinal* research has used methods other than self-report, limiting the ability to assess convergence of trends in self- and informant-reported personality change.

Moreover, a more comprehensive perspective on self–other agreement involving different indicators of stability and change is lacking in the literature. In fact, even if self- and informant-reported mean-level trajectories converged, this would not mean that dyads agree in their ratings of an individual's trajectory. To address this question, research on self–other agreement in individuals' personality change trajectories is needed.

### *Self–other personality agreement*

Self–other personality agreement can be defined as the convergence of self- and informant-reports of an individual's personality traits (e.g. Funder et al., 1995; Kurtz & Sherker, 2003; Watson et al., 2000). A large body of cross-sectional research has found medium to large effect sizes for estimates of self–other agreement in broad personality traits such as the Big Five (Beer

et al., 2013; Biesanz & West, 2004; Funder et al., 1995; McCrae et al., 1998; Vazire, 2010). For example, in a cross-sectional study of over 10,000 Facebook users, Rohrer et al. found correlations between self- and informant-reports ranging from  $r = .30$  for conscientiousness to  $r = .41$  for extraversion (Rohrer et al., 2018). Watson et al. (2000) also found significant cross-sectional self–other agreement in personality ratings across friendship dyads, dating couples, and married couples with effect sizes ranging from  $r = .34$  for agreeableness in friendship dyads to  $r = .61$  for extraversion in married dyads. Moreover, a meta-analysis of 36 studies on cross-sectional self–other agreement found strong convergence (ranging from  $r = .46$  for agreeableness to  $r = .62$  for extraversion) between self- and informant-reports of the Big Five (Connolly et al., 2007). In summary, cross-sectional research provided convincing evidence that self- and informant-reports exhibit moderate agreement in personality trait ratings. However, the cross-sectional nature of this research precludes conclusions regarding self–other agreement about personality change. Longitudinal studies are needed to address questions concerning the degree to which others agree with self-reports of personality trait change over time.

Only few longitudinal studies have examined self–other agreement in personality traits across time. Most of these studies focused on personality development in children and adolescents (Branje et al., 2007<sup>1</sup>; Göllner et al., 2017; Luan et al., 2017). To the best of our knowledge, only three studies examined agreement in personality ratings in adult samples, of which only two examined self–other agreement in personality *change* as we do here. These studies provided mixed evidence concerning the level of self–other agreement in personality change.

First, Kurtz and Sherker (2003) examined self–other agreement in 206 newly acquainted roommates who provided ratings of their own and their roommate’s Big Five personality traits at 2 and 15 weeks into living together. They expected and found an increase in self–other agreement in ratings of openness and agreeableness across the 13-week study interval. Although this research provided evidence that people who have known each other longer tended to agree more in their personality ratings, it lacked an analysis of self–other agreement about personality change.

Second, Watson and Humrichouse (2006) assessed 301 newly married couples twice across a 2-year interval to examine similarity in Big Five personality stability and change between methods of report as well as self–other agreement of personality change among spouses. They found the self- and spousal-reported

rank-order stabilities to be quite similar, with an average rank-order of  $r = .78$  in self-reports and  $r = .77$  in spousal-reports. In contrast to the large self-report literature on personality maturation (Bleidorn, 2015; Specht et al., 2014), they found decreases in most spousal-rated traits during the first 2 years of marriage. Specifically, spousal-reports but not self-reports indicated significant mean-level decreases in extraversion, openness, agreeableness, and conscientiousness, and no significant change in neuroticism. Notably, not all couples showed these negative trends as indicated by significant individual differences in change. This study also examined the agreement between self- and spousal-reported change in personality traits. Results indicated significant, albeit low, correlations between self- and spousal-reported change in conscientiousness ( $r = .20$ ) and neuroticism ( $r = .19$ ), but negligible, insignificant correlations for change in openness ( $r = -.01$ ), agreeableness ( $r = .04$ ), and extraversion ( $r = .08$ ). These findings imply that, while spousal reports of personality traits tend to line up well with baseline self-reports at the beginning of a marriage, they may provide a different perspective on personality change over the course of a marriage.

The above reviewed studies made important contributions to the literature on self–other agreement. However, each of these studies had specific constraints that limited the conclusions that can be drawn about personality trait development. In particular, Kurtz and Sherker’s study (2003) spanned a rather short time span of only 13 weeks which precluded an examination of self–other agreement in personality *change*. Watson and Humrichouse (2006), on the other hand, collected personality data over 2 years; however, with only two waves of data, they were not able to model change parameters using more reliable approaches as latent growth curve modeling (LGM; cf. Willett & Sayer, 1994).

In a recent study, Oltmanns et al. (2019) addressed this issue and used LGM to compare the mean-level trajectories of self- and informant-reported Big Five traits as well as self–other agreement in personality trait change in a sample of older adults ( $M$  age = 62.5 years at first assessment). Participants provided three waves of Big Five self- and informant-reports (significant others, family, or friends) over a period of 6.5 years. Consistent with other research on older adults, this study found evidence for mean-level changes in some but not all Big Five traits (e.g. Kandler et al., 2015; Roberts et al., 2006; Schwaba & Bleidorn, 2018). Specifically, self-report ratings indicated significant decreases in extraversion and openness, and no significant changes in the other trait domains. Combined informant-reports of significant others, family, and/or friends also indicated significant decreases in extraversion and no change in emotional stability; however, they additionally suggested decreases in agreeableness, and

<sup>1</sup>Branje et al. (2007) studied self- and other-reported personality development in 285 Dutch families across 3 years, thus incorporating both adolescents and adults.

conscientiousness, and no change in openness, challenging some seemingly established findings that are based on self-report research only.

This study also investigated self–other agreement in personality trait change. They found evidence for some agreement in self- and informant-reported Big Five trait change. Specifically, the correlations between self- and informant-reports of change in personality traits ranged from  $r = .53$  for agreeableness to  $r = .88$  for neuroticism, indicating substantial agreement in self–other reports of personality change over the study period. These correlations are in sharp contrast with the small to null correlations found by Watson and Humrichouse (2006). These differences may partly reflect differences in study design, particularly the average age of participants, timing of assessments, and methods of report. Overall, Oltmanns et al. (2019) demonstrated that there can be self–other agreement in personality change despite differences between normative self- and informant-reported Big Five trajectories. However, more research is needed to examine when and to what degree self- and informant-reports of personality change converge versus diverge.

In summary, longitudinal research on self–other agreement extends cross-sectional research by assessing agreement in reports of personality trait change. Although the few existing studies on this topic yielded mixed and sometimes conflicting results, an important finding to emerge from all studies is significant differences in self–other agreement between dyads. This finding naturally leads to the question whether there are certain characteristics of dyads that enhance or constrain convergence in reports of personality change. Identifying relevant moderators that shape the degree to which self and informants agree may help in illuminating these differences.

*Moderators of convergence of self- and informant-reports of personality change.* Past research consistently found substantial differences in self–other agreement across studies as well as within studies across dyads. Theory and past research have converged on at least three important sources of this variation: (1) the nature and quality of the dyadic relationship, (2) the visibility of the trait being assessed, and (3) the rater (Connolly et al., 2007; Funder, 1995, 2012). Given the lack of existing research on longitudinal agreement in personality change, the following review of moderators of self–other agreement primarily focuses on findings derived from cross-sectional studies.

The nature and quality of the relationship appears to be associated with the similarity between self- and informant-reported trait trajectories as well as the level of self–other agreement. For example, Watson and Humrichouse's (2006) study of newlyweds suggested that individual differences in marital satisfaction shaped the way in which spouses described the personalities of their partners. Specifically, those

couples who reported more pronounced decreases in marital satisfaction over a 2-year period also reported more negative trait changes. Similarly for self–other agreement, a meta-analysis of cross-sectional studies on self–other agreement found that people judged their partners more positively when they reported higher levels of relationship quality (Fletcher and Kerr, 2010).

Interpersonal factors such as relationship length and closeness have also been associated with the amount and quality of information that informants have access to when judging others' personality traits. Typically, intimate dyads who tend to spend more time together should have access to a larger amount of personality-relevant information. A large body of cross-sectional research has supported this position, indicating that level of acquaintance moderates self–other agreement in personality traits (Biesanz et al., 2007; Connolly et al., 2007; Funder et al., 1995). For example, Watson et al. (2000) found a medium effect size for self–other agreement in friendship dyads, a slightly larger effect in dating couples, and an even larger effect size in married couples. This finding has often been referred to as *acquaintanceship effect*. Consistent with cross-sectional research, longitudinal data from Kurtz and Sherker (2003) indicated increases in agreement between roommates for openness and agreeableness with increased acquaintance after a 13-week period. However, it should be noted that substantial increases in self–other agreement likely occur within the first few months of a relationship (Allik et al., 2016; Human et al., 2020), and time spent in the relationship thereafter may not significantly add to the level of self–other agreement.

Theory and cross-sectional research have further indicated that trait visibility—the degree to which trait-relevant behavior can be readily observed by an outside informant—may moderate self–other agreement. Extraversion and agreeableness are viewed as particularly visible traits due to the prevalence of observable behavior associated with these traits. In contrast, emotional stability is usually considered the least visible trait of the Big Five, particularly among mere acquaintances, because of its high affective load. Conscientiousness and openness are thought to be moderately visible (Wilt & Revelle, 2015). In other words, traits that are characterized by more behavioral content should be more visible whereas traits made up of more affective or cognitive content should be less visible. Consistent with this hypothesis, research has consistently found self–other agreement to be highest for extraversion (Beer & Watson, 2008b; Beer et al., 2013; Connolly et al., 2007; John & Robins, 1993). In contrast, Watson et al. (2000) found the lowest agreement for emotional stability between friends and dating couples; although agreement was higher in married couples, suggesting that more trait-relevant behavior about negative emotions is revealed in the privacy of close



relationships than in public. However, despite its relatively clear behavioral associations, the trait with the lowest self–other agreement tends to be agreeableness (Beer et al., 2013; Connolly et al., 2007; John & Robins, 1993; Watson et al., 2000). There are at least two possible explanations for the comparatively lower level of self–other agreement in agreeableness. First, others may misinterpret the underlying intentions of overt behavior. For example, providing critical feedback may be well intended but interpreted as confrontational and unpleasant by outside observers. Second, others reporting on a target’s level of agreeableness may be reluctant to report low levels due to the negative connotation associated with low levels of agreeableness (Rau et al., 2019).

Considering both acquaintanceship and trait visibility effects, existing findings follow a coherent pattern: Ratings by strangers show little agreement and agreement is often restricted to extraversion (Beer & Watson, 2008b; Vazire, 2010); self–other agreement tends to be somewhat higher in ratings by recently acquainted individuals, and even higher in ratings of close peers (Beer & Watson, 2008a, 2008b; Funder et al., 1995; Hayes & Dunning, 1997) and dating couples (Watson et al., 2000). Yet, only ratings by spouses tend to show substantial self–other agreement in each of the Big Five traits, ranging from medium effect sizes for agreeableness to large effect sizes for extraversion (Beer et al., 2013; Watson et al., 2000).

### *The present study*

The goal of the present longitudinal study was to examine personality stability and change in both self- and partner-reports and self–other personality agreement across four assessments in a sample of 255 married and cohabitating Dutch couples ( $N = 510$ ; age: 19–45, <https://osf.io/c3hb4/>, 2018). This research expands on the existing literature on personality development in two important ways.

First, we used partner-report data to examine the convergent validity of self- and partner-reported Big Five personality trait development in early adulthood. We did this by comparing self- and partner-reported Big Five rank-order stabilities and mean-level trajectories. Consistent with past research, we expected to find significant and high rank-order stabilities for all Big Five traits as well as mean-level increases in emotional stability, agreeableness, and conscientiousness in both self- and partner-reports.

Second, we examined self–other agreement in self- and partner-reports of personality concurrently and across time. Consistent with evidence for cross-sectional self–other personality agreement in romantic couples, we hypothesized that cross-sectional agreement would be moderately high (Watson et al., 2000; Watson & Humrichouse, 2006). To model agreement in change, we estimated bivariate LGMs of self- and partner-reported traits for each of the Big

Five traits. Self–other agreement in personality change may result from at least two mechanisms that are not mutually exclusive. First, self–other personality agreement may reflect a spouse’s perception of actual changes in their partner’s typical pattern of thoughts, feelings, and behaviors. Second, independent of whether actual change has occurred or not, spouses may develop shared beliefs about their personality change. Given that few studies have examined self–other agreement over time, we adopted an exploratory approach in examining self–other agreement in personality trait change.

We expected to find significant differences in self–other agreement across couples. Consistent with past research, we predicted that interpersonal factors and trait visibility would influence the degree of self–other agreement. Specifically, we predicted that a longer relationship length and cohabitation status would be associated with higher agreement between self- and partner-reports both cross-sectionally and over time. Another interpersonal factor we considered was parenthood status—a role that may reflect a more intimate relationship between partners. This higher level of intimacy could lead to individuals feeling more comfortable to exhibit a range of trait-relevant behavior. Thus, we hypothesized higher agreement in both cross-sectional and longitudinal ratings of self- and partner-reports from parents compared to dyads without children. Overall, we expected substantial self–other agreement for each of the Big Five traits. Considering the visibility of the trait and considering prior research (Connolly et al., 2007; Watson et al., 2000), however, we expected relatively higher agreement in extraversion, moderate agreement in conscientiousness and openness, and lower agreement in emotional stability and agreeableness. Finally, we explored gender differences in self–other agreement (Chan et al., 2011) by testing whether agreement was different depending on whether women or men provided other-reports of their partners’ personality traits.

## **Method**

### *Sample and procedure*

Data came from a four-wave longitudinal study of first-time parents and couples without children. For a detailed description of all measures collected as part of this project, see [osf.io/c3hb4/](https://osf.io/c3hb4/). We did not preregister our aforementioned hypotheses. Previous published work with this data includes Van Scheppingen et al. (2018), <https://osf.io/xqpcu/>. This study overlaps with the present submission only with regards to the sample assessed. For supplementary material, data, and its corresponding reproducible code, see <https://osf.io/z86ku/>.

Both subsamples completed four online assessments across the 20-month study period. Parents

**Table 1.** Descriptives including Cronbach's alpha, raw means, and standard deviations of self- and partner-reported personality traits.

|    | Wave 1     |          | Wave 2     |          |            |          | Wave 3     |          |            |          | Wave 4     |          |            |          |            |     |
|----|------------|----------|------------|----------|------------|----------|------------|----------|------------|----------|------------|----------|------------|----------|------------|-----|
|    | Self       | Partner  | Self       | Partner  | Self       | Partner  | Self       | Partner  | Self       | Partner  | Self       | Partner  |            |          |            |     |
|    | M (SD)     | $\alpha$ | M (SD)     | $\alpha$ | M (SD)     | $\alpha$ | M (SD)     | $\alpha$ | M (SD)     | $\alpha$ | M (SD)     | $\alpha$ | M (SD)     | $\alpha$ |            |     |
| ES | 3.24(0.63) | .84      | 3.29(0.80) | .89      | 3.29(0.64) | .85      | 3.27(0.82) | .89      | 3.29(0.67) | .87      | 3.28(0.78) | .88      | 3.28(0.66) | .86      | 3.23(0.83) | .90 |
| E  | 3.55(0.60) | .83      | 3.72(0.65) | .83      | 3.52(0.63) | .85      | 3.71(0.61) | .80      | 3.51(0.60) | .84      | 3.65(0.61) | .82      | 3.50(0.60) | .83      | 3.62(0.61) | .82 |
| O  | 3.39(0.53) | .76      | 3.29(0.58) | .77      | 3.40(0.56) | .77      | 3.30(0.61) | .78      | 3.40(0.57) | .79      | 3.31(0.62) | .79      | 3.37(0.57) | .78      | 3.29(0.66) | .81 |
| A  | 3.67(0.46) | .71      | 3.81(0.53) | .76      | 3.67(0.47) | .71      | 3.82(0.55) | .79      | 3.71(0.46) | .71      | 3.82(0.55) | .79      | 3.70(0.44) | .68      | 3.80(0.60) | .82 |
| C  | 3.56(0.53) | .79      | 3.67(0.65) | .82      | 3.58(0.54) | .80      | 3.67(0.65) | .83      | 3.58(0.52) | .79      | 3.69(0.60) | .81      | 3.57(0.52) | .78      | 3.70(0.65) | .84 |

A: agreeableness; C: conscientiousness; E: extraversion; ES: emotional stability; O: openness; Partner: partner report; Self: self-report.

completed the first assessment 6 weeks before the due date and completed the remaining three assessments when the child was 6, 12, and 18 months old. Nonparents completed the assessments according to the same assessment schedule.

In the present study, we used data from all participants who provided self- and partner-reports of personality for at least one assessment wave. The total sample consisted of 248 parents and 294 nonparents, ranging from 19 to 45 years of age (52.5% female,  $M$  age = 27.01 years,  $SD$  = 4.61 years). Of this sample, 42 individuals participated without their partner, resulting in a total of 255 complete dyads ( $N$  = 510). The majority of participants in the parent subsample were married (55.9%) or cohabitating (43.3%), whereas most participants in the nonparent sample were cohabitating (54.5%) or not living together (35.7%). As expected, parents were indeed more likely to live together ( $r$  = .26) and tended to be in longer relationships ( $r$  = .30). Across the four assessment waves, the sample sizes of complete dyads were 236, 176, 170, and 146.

## Measures

**Personality.** At each assessment wave, participants reported on their own and their partner's personality traits using the 44-item Big Five Inventory (BFI) Self- and Partner-Report (John et al., 1991). Responses were measured on a scale of 1 (*completely disagree*) to 5 (*completely agree*). Internal consistency, as shown in Table 1, was above  $\alpha$  = .70 for all traits and waves, except for agreeableness ( $\alpha$  = .68 at wave 4 for self-report).

**Relationship variables.** We used three variables to indicate relationship intimacy: parenthood status, relationship length, and cohabitation status. As described above, parenthood status was recorded during the recruitment stage of this study.<sup>2</sup> Relationship length was measured once at Wave 4. Participants were asked "How long have you been in a relationship with your current partner?" and reported the number of years and months. Length ranged from 2 to 25 years with the average length

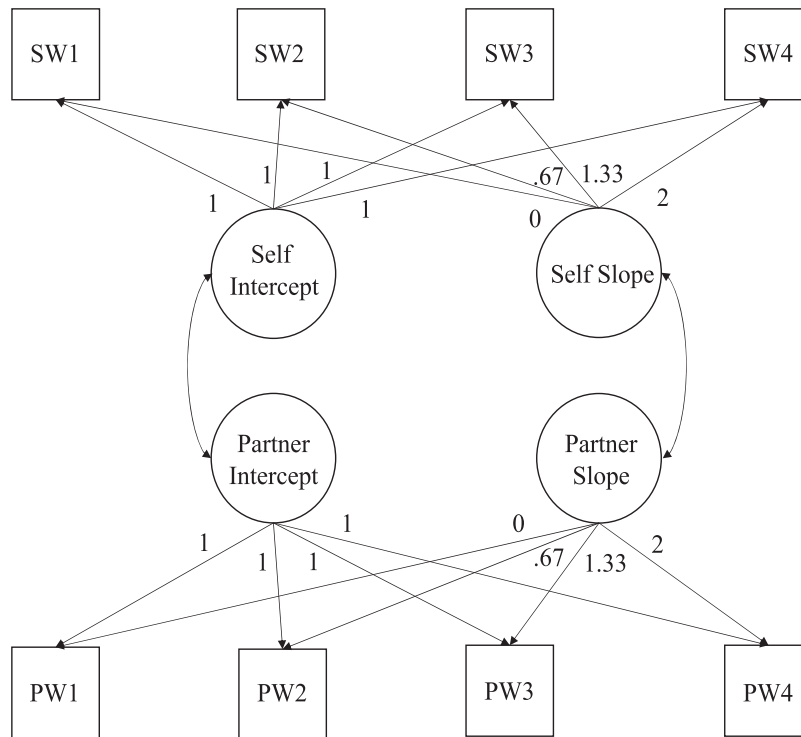
of couples' romantic relationship being 6.81 years ( $SD$  = 3.75) and the median being 6 years. Cohabitation was measured once at Wave 1. Participants who reported "married," "living together," or "cohabitation agreement" were scored as cohabitating ( $N$  = 375 at wave 1); participants who reported "other" were scored as not cohabitating ( $N$  = 116 at Wave 1).

## Analyses

We analyzed all data using the lavaan package version 0.6-3 in R version 3.5.1 (R Core Team, 2013; Rosseel, 2012). We used Full Information Maximum Likelihood (FIML) estimation to account for missing data. FIML estimation models use all available data from all participants. To evaluate absolute model fit, we used the Comparative Fit Index (CFI) and the Root Mean Squared Error of Approximation (RMSEA; Hu & Bentler, 1998, 1999). We compared fits of nested models using  $\chi^2$  difference tests.

A common concern associated with studying dyads is the possible dependency in the data (Kashy & Snyder, 1995; Kenny, 1995). In other words, members within a dyad may relate to each other in a systematic way that influences the data. However, consistent with past research, we found little evidence for similarity between partners' personality traits with the exception of openness to experience (Beer et al., 2013; Watson et al., 2000, 2004).<sup>3</sup> Considering the nature of the present sample, we interpret self-other agreement data in the context of romantic couples. More longitudinal multi-rater research is needed to examine the degree to which these results generalize to other types of dyads (Branje et al., 2007; Luan et al., 2017).

**Convergence of rank-order stabilities and mean-level change.** We first examined similarities in rank-order stabilities of self- and partner-reported Big Five traits using autoregressive models. In these models, the trait scores at each wave were regressed on the trait score assessed at the previous wave across the four waves. We then used model comparison



**Figure 1.** Bivariate latent growth curve model of self- and partner-reported Big Five traits. Self–other agreement is indicated by the double-headed arrow from Self Slope to Partner Slope. W1, W2, W3, and W4 indicate the observed scores at the four assessment waves for each trait used to estimate the latent intercept and latent slope factors, with S and P denoting self- or partner-report.

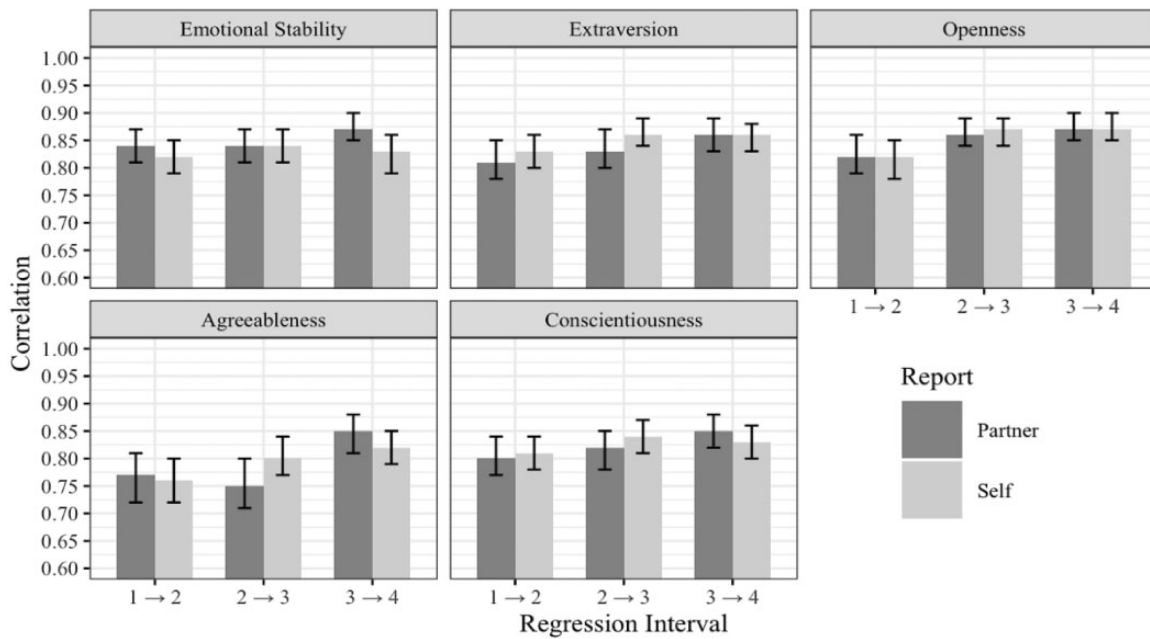
strategies to test whether the rank-order stability coefficients could be constrained to be equal across assessment intervals without a significant decrease in model fit. If this was not the case, we estimated a series of unconstrained models to identify the interval(s) for which the rank-order stabilities could not be constrained to be equal. Critical for our research question, we then tested whether rank-order stability coefficients could also be constrained to be equal across self- and partner-reports.

To examine the degree of similarity in mean-level trajectories of self- and partner-reported personality traits, we estimated univariate LGM for each of the self- and partner-reported Big Five traits (Duncan et al., 2006). As shown in Figure 1, the factor loadings for determining the slope were fixed with respect to the 6-month assessment intervals over the study period. We then plotted the model-implied Big Five trajectories for both methods of report to visualize their similarity. Next, we combined the univariate LGM of self- and partner-reports into a bivariate LGM for each of the five traits and tested whether the self- and partner-report slope parameters (means and variances) could be constrained to be equal without a significant decrease in model fit (relative to the unconstrained model), indicating no significant difference in self- and partner-reported personality trait

trajectories. We allowed intercept–slope correlations in all models.

**Self–other agreement.** We examined cross-sectional agreement between self- and partner-reports using Pearson’s *r* correlations for each trait at each assessment wave (5 traits × 4 assessment waves = 20 correlations). To examine longitudinal self–other agreement in personality change, we again used bivariate LGM of self- and partner-reports and tested whether there was correlated change between the latent slope factors of self- and partner-reports (Hertzog et al., 2006). A significant slope–slope correlation would indicate that self-reported change in a personality trait is related to partner-reported change in that trait.

We used a multiple-group structural equation modeling (SEM) framework to analyze the moderating effects of interpersonal factors on cross-sectional self–other agreement, and then a multiple-group bivariate LGM when analyzing moderating effects on longitudinal self–other agreement. First, we split participants into two groups for each moderator: parents (*N* = 241) and nonparents (*N* = 291), short (≤ 6 years, *N* = 194) and long relationships (> 6 years, *N* = 146),<sup>4</sup> participants cohabitating (*N* = 375) and not cohabitating (*N* = 116), and men (*N* = 236) and women (*N* = 255).



**Figure 2.** Self- and partner-reported rank-order stabilities for Big Five personality traits across three 6-month intervals.

## Results

### Attrition analyses

To examine the potential effects of attrition, we compared participants who dropped out after the first assessment wave to participants who completed more than one wave of personality data with regard to age, gender, and personality. Participants who dropped out after wave 1 had significantly lower scores in emotional stability ( $M_{(\text{dropout})} = 3.13$ ,  $SD = 0.61$ ;  $M_{(\text{continue})} = 3.27$ ,  $SD = 0.63$ ;  $t(182.36) = 2.26$ ,  $p < .05$ ) and agreeableness ( $M_{(\text{dropout})} = 3.53$ ,  $SD = 0.45$ ;  $M_{(\text{continue})} = 3.70$ ,  $SD = 0.46$ ;  $t(179.17) = 3.67$ ,  $p < .001$ ) than participants who did not drop out. Dropouts and continuing participants did not differ in other traits, age, or gender.

### Rank-order stability

Rank-order stabilities were high ( $> .70$ ) across both assessment waves and methods of report (Figure 2), ranging from .76 to .87 for self-reports and .75 to .87 for partner-reports. After constraining rank-order stabilities to be equal across time, constraining the parameters of self- and partner-reported stability coefficients to be equal did not decrease the fit of the models for emotional stability ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 10.41/5$ ,  $p = .06$ ), extraversion ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 6.67/5$ ,  $p = .25$ ), openness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 3.93/5$ ,  $p = .56$ ), and conscientiousness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 10.75/5$ ,  $p = .06$ ). However, constraining self- and partner-reported stabilities for agreeableness to be equal across the third interval (from wave 3 to wave 4) resulted in a significant decrease in model fit ( $\Delta\text{CFI} = .004$ ,  $\Delta\chi^2/\Delta\text{df} = 13.77/5$ ,  $p = .02$ ), indicating

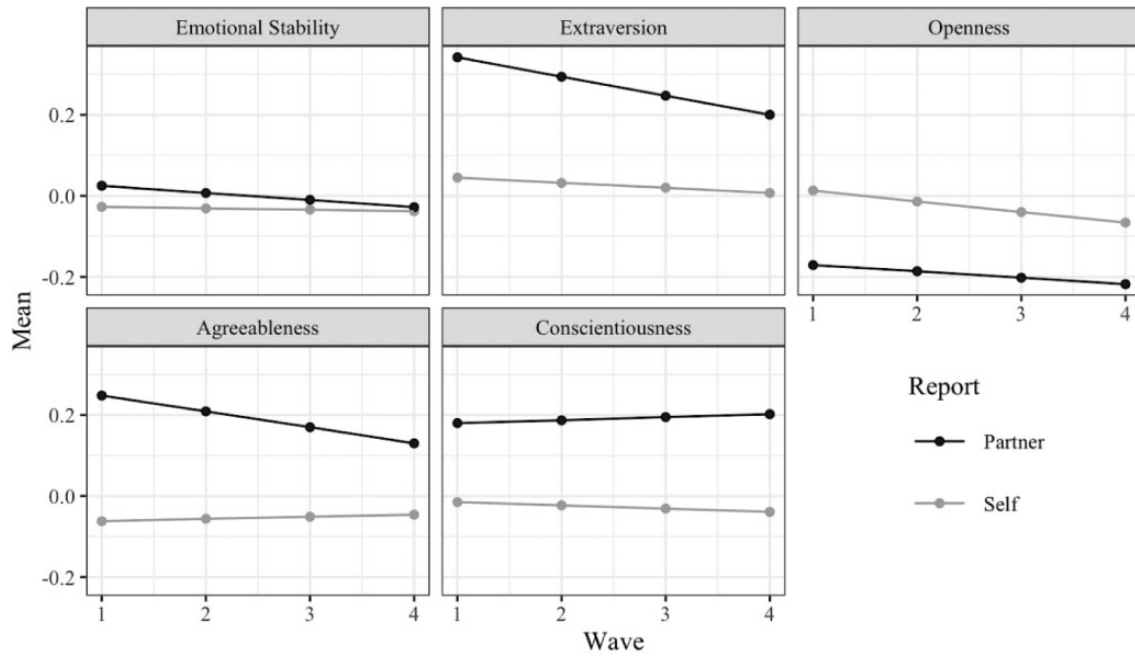
that the rank-order stability of self-reports was lower than of partner-reports during the third interval. Overall, with only one exception, results indicated highly similar rank-order stabilities of self- and partner-reports across the three 6-month intervals. Correlations derived from these models are presented in Table S2 of the online supplementary material.

### Mean-level change

Figure 3 depicts the model-implied mean-level trajectories of self- and partner-reported Big Five traits across the study period derived from the univariate LGMs described above (for the raw personality trajectory scores, see Table 1; for the model-implied personality trajectory scores, see Table S3 of the online supplementary material). Both self- and partner-reported slopes indicated little to no mean-level change in all Big Five traits over the four assessment waves (see Table 2). Specifically, self-reported slopes indicated small mean-level decreases in openness ( $M = -0.02$ ,  $p = .01$ , 95% CI  $[-0.04, -0.01]$ ), and partner-reported slopes indicated small mean-level decreases in extraversion ( $M = -0.04$ ,  $p < .001$ , 95% CI  $[-0.07, -0.02]$ ), and agreeableness ( $M = -0.03$ ,  $p = .02$ , 95% CI  $[-0.05, -0.04]$ ). Moreover, the variance for the slopes in extraversion, openness, and conscientiousness in self-reports and openness and conscientiousness in partner-reports were not significant.

After combining the univariate LGM for each method report, the bivariate LGM for emotional stability, extraversion, openness, agreeableness, and conscientiousness fit the data well (all CFI  $\geq .99$ , RMSEA  $< .05$ ). Constraining the self- and partner-





**Figure 3.** Trajectories of self- and partner-reported Big Five personality traits across four assessment waves (values are z-standardized).

slope parameters to be equal did not decrease the fit of the models for emotional stability ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 2.80/2$ ,  $p = .25$ ), openness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 1.01/2$ ,  $p = .60$ ), and conscientiousness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.86/2$ ,  $p = .65$ ), suggesting no difference between self- and partner-reported trait trajectories of these traits. However, constraining self- and partner-slope parameters for extraversion ( $\Delta\text{CFI} = .002$ ,  $\Delta\chi^2/\Delta\text{df} = 8.72/2$ ,  $p = .01$ ) and agreeableness ( $\Delta\text{CFI} = .003$ ,  $\Delta\chi^2/\Delta\text{df} = 7.59/2$ ,  $p = .02$ ) resulted in a significant decrease in model fits. Specifically, partner-reports indicated a more negative slope for extraversion than did self-reports. Moreover, the slope for partner-reported agreeableness was negative whereas the slope for self-reported agreeableness was slightly positive. Overall, these results indicate that self- and partner-reports of mean-level change were largely consistent, with exceptions for extraversion and agreeableness.

### Cross-sectional self–partner agreement

Table 3 shows the correlations between self- and partner-reported personality traits for each assessment wave. Consistent with our hypothesis and in line with past research on self–other agreement in romantic dyads, cross-sectional agreement was strong, ranging from  $r = .43$  for agreeableness to  $r = .63$  for emotional stability, with extraversion just below at  $r = .61$ .

**Moderators of cross-sectional agreement.** We first tested the effect of parenthood status on agreement and

found that constraining these groups to be equal did not decrease the fit of the models for emotional stability ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.39/1$ ,  $p = .53$ ), extraversion ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.86/1$ ,  $p = .35$ ), openness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.09/1$ ,  $p = .77$ ), agreeableness ( $\Delta\text{CFI} = .001$ ,  $\Delta\chi^2/\Delta\text{df} = 3.62/1$ ,  $p = .06$ ), and conscientiousness ( $\Delta\text{CFI} = .001$ ,  $\Delta\chi^2/\Delta\text{df} = 2.44/1$ ,  $p = .12$ ), suggesting no differences between parents and nonparents in regards to cross-sectional self–other agreement. Parent and nonparent correlations between self- and partner-reported personality traits for each assessment wave are presented in Table S4 of the online supplementary material.

Likewise, constraining agreement between participants in relatively shorter and longer relationships to be equal did not decrease the fit of the models for emotional stability ( $\Delta\text{CFI} = .001$ ,  $\Delta\chi^2/\Delta\text{df} = 1.31/1$ ,  $p = .25$ ), extraversion ( $\Delta\text{CFI} = .001$ ,  $\Delta\chi^2/\Delta\text{df} = 3.25/1$ ,  $p = .07$ ), openness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.50/1$ ,  $p = .48$ ), and agreeableness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.16/1$ ,  $p = .69$ ), suggesting no difference between these groups. However, constraining the group parameters to be equal for conscientiousness did produce a significantly worse fitting model ( $\Delta\text{CFI} = .001$ ,  $\Delta\chi^2/\Delta\text{df} = 4.06/1$ ,  $p = .04$ ), in which couples in shorter relationships ( $\leq 6$  years,  $r = .56$ ) had higher cross-sectional self–other agreement than couples in longer relationships ( $> 6$  years,  $r = .51$ ) in this trait.

Constraining agreement between participants cohabitating and not cohabitating to be equal did not decrease the fit of the models for emotional stability ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.14/1$ ,  $p = .70$ ),

**Table 2.** Univariate models of change.

|                | Intercept |              |          |            | Slope    |              |          |          |             |          |
|----------------|-----------|--------------|----------|------------|----------|--------------|----------|----------|-------------|----------|
|                | Mean      |              | Variance |            | Mean     |              |          | Variance |             |          |
|                | Estimate  | 95% CI       | Estimate | 95% CI     | Estimate | 95% CI       | <i>p</i> | Estimate | 95% CI      | <i>p</i> |
| <b>Self</b>    |           |              |          |            |          |              |          |          |             |          |
| ES             | 3.25      | [3.20, 3.31] | 0.34     | [.29, .38] | -0.00    | [-.03, .02]  | .74      | 0.01     | [.00, .02]  | .02      |
| E              | 3.55      | [3.50, 3.60] | 0.31     | [.26, .35] | -0.01    | [-.03, .01]  | .17      | 0.00     | [-.01, .01] | .81      |
| O              | 3.40      | [3.35, 3.44] | 0.23     | [.20, .27] | -0.02    | [-.04, -.01] | .01      | 0.00     | [-.00, .01] | .33      |
| A              | 3.66      | [3.62, 3.70] | 0.17     | [.14, .19] | 0.00     | [-.01, .02]  | .66      | 0.01     | [-.00, .01] | .02      |
| C              | 3.56      | [3.52, 3.61] | 0.23     | [.20, .26] | -0.01    | [-.02, .01]  | .46      | 0.00     | [-.00, .01] | .24      |
| <b>Partner</b> |           |              |          |            |          |              |          |          |             |          |
| ES             | 3.29      | [3.22, 3.36] | 0.54     | [.46, .62] | -0.02    | [-.04, .01]  | .23      | 0.02     | [.01, .04]  | .002     |
| E              | 3.73      | [3.67, 3.78] | 0.34     | [.28, .39] | -0.04    | [-.07, -.02] | < .001   | 0.01     | [.00, .02]  | .01      |
| O              | 3.29      | [3.25, 3.35] | 0.28     | [.23, .32] | -0.01    | [-.03, .01]  | .21      | 0.01     | [-.00, .02] | .06      |
| A              | 3.80      | [3.76, 3.85] | 0.22     | [.18, .25] | -0.03    | [-.05, -.04] | .02      | 0.02     | [.01, .03]  | < .001   |
| C              | 3.67      | [3.62, 3.73] | 0.32     | [.27, .38] | 0.01     | [-.02, .03]  | .61      | 0.01     | [-.00, .02] | .19      |

All intercept values:  $p < .001$ .

A: agreeableness; C: conscientiousness; E: extraversion; ES: emotional stability; O: openness.

**Table 3.** Pearson's *r* cross-sectional self–other agreement correlations.

|        | Emotional stability | Extraversion | Openness | Agreeableness | Conscientiousness |
|--------|---------------------|--------------|----------|---------------|-------------------|
| Wave 1 | .63                 | .58          | .61      | .45           | .57               |
| Wave 2 | .59                 | .61          | .60      | .45           | .51               |
| Wave 3 | .63                 | .59          | .62      | .47           | .51               |
| Wave 4 | .63                 | .60          | .58      | .43           | .55               |

$p < .001$  for all values.

extraversion ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.39/1$ ,  $p = .53$ ), openness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.62/1$ ,  $p = .43$ ), agreeableness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 1.26/1$ ,  $p = .26$ ), and conscientiousness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 2.26/1$ ,  $p = .13$ ), suggesting no difference between these groups. Neither did constraining agreement between men and women to be equal for emotional stability ( $\Delta\text{CFI} = .001$ ,  $\Delta\chi^2/\Delta\text{df} = 2.51/1$ ,  $p = .11$ ), extraversion ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 1.67/1$ ,  $p = .20$ ), openness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.11/1$ ,  $p = .74$ ), agreeableness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.27/1$ ,  $p = .60$ ), and conscientiousness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.01/1$ ,  $p = .89$ ), suggesting no difference between these groups.

In summary, we only found one significant difference for relationship length in levels of cross-sectional agreement for conscientiousness. However, this difference was inconsistent with our predictions suggesting that a shorter relationship length was associated with higher levels of agreement.

### Self–other agreement in personality change

The results of the bivariate LGM are displayed in Table 4. Bivariate models for extraversion, openness, and conscientiousness did not converge due to the lack of significant variance in the slopes (see

Table 2). As such, we only examined the correlation of self- and partner-reported slopes in emotional stability and agreeableness. Results indicated significant self–other agreement for change in both emotional stability ( $r = .51$ ,  $p = .001$ ) and agreeableness ( $r = .36$ ,  $p = .042$ ).

### Moderators of self–other agreement in personality change.

We first tested a multiple-group bivariate LGM to examine the difference in self–other agreement between parents and nonparents. Constraining the slope–slope correlations across parents and nonparents did not lead to a significant decrease in model fit for emotional stability ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.04/1$ ,  $p = .85$ ) and agreeableness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 2.85/1$ ,  $p = .09$ ) suggesting no differences between these groups. Likewise, constraining the slope–slope across participants in shorter ( $\leq 6$  years) and longer ( $> 6$  years) relationships did not lead to a significant decrease in model fit for emotional stability ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.04/1$ ,  $p = .83$ ) or agreeableness ( $\Delta\text{CFI} = .001$ ,  $\Delta\chi^2/\Delta\text{df} = 0.29/1$ ,  $p = .59$ ). Neither did constraining the slope–slope correlations to be equal across participants cohabitating and not cohabitating for emotional stability ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.88/1$ ,  $p = .35$ ) or agreeableness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.48/1$ ,  $p = .49$ ). We also found no

**Table 4.** Bivariate models of change.

|                     | Self-partner intercept correlation |          |            |          | Self-partner slope correlation |          |             |          |
|---------------------|------------------------------------|----------|------------|----------|--------------------------------|----------|-------------|----------|
|                     | Covariance (SE)                    | <i>r</i> | 95% CI     | <i>p</i> | Covariance (SE)                | <i>r</i> | 95% CI      | <i>p</i> |
| Emotional stability | 0.304 (.024)                       | .72      | .663, .778 | < .001   | 0.009 (.003)                   | .51      | .163, .865  | .001     |
| Agreeableness       | 0.110 (.011)                       | .58      | .501, .663 | < .001   | 0.004 (.002)                   | .36      | -.006, .734 | .042     |

significant differences between men and women. Specifically, constraining the slope-slope correlations to be equal across models with other-reports by men versus women did not lead to a significant decrease in model fit for emotional stability ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 2.58/1$ ,  $p = .11$ ) or agreeableness ( $\Delta\text{CFI} = .00$ ,  $\Delta\chi^2/\Delta\text{df} = 0.88/1$ ,  $p = .35$ ). In summary, we found no evidence for moderating effects of parenthood status, relationship length, cohabitation, or gender.

## Discussion

In the present study, we tested whether the trends in personality rank-order stability and mean-level change typically observed in self-report data can be replicated in partner-reports. In addition, we examined the degree to which self- and partner-reports of personality *change* converge. Below, we highlight the main findings and discuss theoretical implications, with a particular emphasis on the importance of multi-method assessment in longitudinal personality research.

### Self- and partner-reported personality

**Rank-order stability.** Results of the present study indicated that rank-order stabilities from partner reports tend to converge with those obtained from self-reports. Consistent with the large body of self-report research on personality rank-order stability (Roberts & DelVecchio, 2000), we found substantive rank-order stabilities across 6-month intervals in both self- and partner-reports, ranging between  $r = .75$  and  $r = .87$ . Moreover, with one exception, we found no significant differences in rank-order stabilities between self- and partner-reports for the 30 stabilities we compared. This finding is consistent with previous results from studies that have used other multi-method assessments of personality across time (Dobewall & Aavik, 2016; Watson & Humrichouse, 2006), indicating that we can obtain the same information regarding the stability of personality traits from partner reports as from self-reports.

**Mean-level change.** We found more mixed evidence when assessing personality mean-level change in self- and partner-reports. The self- and partner-reported mean-level trajectories for emotional stability, openness, and conscientiousness were not significantly different between methods of report. However, we found significant differences in the

self- and partner-reported mean-level trajectories of extraversion and agreeableness. Specifically, whereas self-reported trajectories were flat and indicated no significant mean-level change in these traits, partner-reports indicated significant mean-level decreases in these two domains. This finding is consistent with Watson and Humrichouse (2006) who also reported decreases in partner-reported but not in self-reported extraversion and agreeableness (as well as in openness and conscientiousness) in a sample of newlyweds. However, it is important to note here that partner-reported levels in these traits were initially higher than self-reports and were still higher by the final assessment despite the reported decreases. A possible explanation for these differences between self- and partner-reports may be changes in the relationship and partner perceptions (Rau et al., 2019). Specifically, people may be more likely to present their best selves and/or their partners may have a desire to see them through rose-colored glasses early in the relationship. As time goes on, partners may be more likely to see negative personality characteristics in their spouse which may be related to decreases in relationship satisfaction (Fletcher & Kerr, 2010; Watson & Humrichouse, 2006). That is, with time, the degree of positivity bias (rose-colored glasses) the partner has for their spouse likely decreases (Zimmermann et al., 2018).

Alternatively, it is possible that partners are better suited to pick up decreases in these traits which might otherwise go unnoticed in self-report data. Specifically, it may be more difficult for people to notice and report undesirable changes in their own personality traits than their partner's personality traits. Future research is needed to expand on the present findings in at least two important ways. First, longitudinal studies involving informant-reports across social settings would be a step toward expanding on this research. Incorporating a variety of informant-reports could help address the question of accuracy in personality reports. Different social contexts could evoke different trait-relevant behaviors that only observers in that context have access to. For example, coworkers could notice the drive and ambition of someone trying to get a promotion at work whereas only the spouse observes her stress and worry about her family life upon returning home. Combining informant-reports from various social contexts may thus provide a more comprehensive picture of an individual's personality (Funder

et al., 1995). Second, future research should include questions about how confident informants are in their ratings of personality questionnaire items. To the degree that informants agree in their confidence ratings, these could provide information about items that are more or less suitable for informant-report questionnaires.

In summary, we found similarly high rank-order stabilities in self- and partner-report data for all Big Five traits and some evidence for convergence in self- and partner-reported mean-level change in emotional stability, openness, and conscientiousness. Deviations in the mean-level trajectories of agreeableness and extraversion may indicate that partner reports can provide non-overlapping information to complement self-reports in some domains. Indeed, informants, and romantic partners in particular, may be better suited to observe certain aspects of our personality and report more accurately on areas in which we have blind spots (Gallrein et al., 2013; Vazire, 2010). For example, people may describe themselves as agreeable when they engage in some friendly discourse that stimulates communion, whereas the partner may observe the heated debate and uneasiness of the participant on the other end. However, the accuracy of the different methods of report remains an open question. Future research incorporating a variety of methods of assessment may advance our understanding of possible rater biases in reports of personality stability and change.

### *Self–other agreement*

The second goal of this study was to examine whether and to what degree partner-reports converge with self-reports of personality traits over time. To address this goal, we first examined cross-sectional agreement between self- and partner-reports across the four assessment waves. Three findings stand out.

First, consistent with past research on self–other agreement, we found substantial convergence across all traits at all assessment waves. Second, unlike most studies of self–other agreement (Beer & Watson, 2008b; Connolly et al., 2007), we found the strongest agreement for emotional stability. This result may reflect the closeness of partners in romantic relationships, as past research has also found stronger agreement in emotional stability in samples of romantic couples compared to other dyads (Beer et al., 2013; Watson et al., 2000). Partners in romantic relationships may thus be uniquely suited to rate more ‘private’ traits such as emotional stability, possibly because they have greater access to their partner’s thoughts, feelings, and behaviors. Third, as predicted, self–partner agreement was consistently lower in agreeableness relative to the other Big Five traits across all assessment waves.

There are at least four explanations for the comparatively larger divergence of self- and partner-

reports of agreeableness. First, the slightly lower level of internal consistency of agreeableness relative to the remaining traits may have contributed to the lower level of self–other agreement. Second, the comparatively low level of agreement may arise from inaccuracies in partner-reports. That is, given that agreeableness is a highly evaluative trait, partners may be more reluctant to report disagreeable traits and more motivated to describe their partners as more agreeable (Rau et al., 2019). Third, the lower level of agreement might stem from an inability in partners to acknowledge and interpret behavior relevant to agreeableness (Vazire & Carlson, 2010). Specifically, some behaviors may be very subtle with a hint of malice, however, this nuance may be difficult to pick up by partners. Fourth, both self- and partner-reports could be accurate to some extent but may be reporting on aspects of agreeableness that are not correlated. For example, a behavior that is intended to be sarcastic and cunning could be interpreted as friendly joking by the partner.

We also examined the moderating effects of interpersonal factors and gender on cross-sectional agreement. Contrary to our predictions, we found little to no evidence for moderating effects of these variables. As noted in previous research (Allik et al., 2016; Human et al., 2020), significant increases in self–other agreement likely occur within the first few months of a relationship. Considering that the average relationship length in our sample was 6.81 years, it is likely that most couples had already reached their maximum level of self–other agreement.

We further expected but did not find higher self–other agreement in parents compared to nonparents, suggesting that parents report on each other’s personality at a level of convergence similar to romantic partners without children. Similarly, we did not find evidence for differences between participants who were cohabiting or not, nor across men and women. Overall, our results provided little evidence for the hypothesis that relationship closeness may contribute to higher self–other personality in long-term romantic relationships.

A key question of this study was whether partners agree in their reports of personality change. Consistent with Oltmanns et al. (2019), our findings suggest dyads have high levels of self–other agreement on longitudinal personality change, contradictory to Watson and Humrichouse (2006) who found little evidence for self–other agreement on personality change. We focused on emotional stability and agreeableness—the two traits that exhibited significant individual differences in change across the four-wave research period—and tested whether self-reported changes in these traits were correlated with changes in partner-reports (cf. Oltmanns et al., 2019). As with our results for cross-sectional agreement, we found the highest level of self–other agreement for emotional stability, and a moderately high level of agreement



for agreeableness, implying that changes in emotional stability may be easier to detect, especially for romantic dyads. It is possible that changes in less observable traits may be more noticeable to others due to discrepancies in broad demeanor as opposed to highly observable traits where changes may be more subtle and harder to detect. Future research is needed to further examine the relationship between trait visibility and agreement in change.

Overall, the observed self–other agreement in change suggests that, to the degree that individuals report significant changes in their personality traits that deviate from the normative trends, partners are able to note and report these changes. This agreement in personality change may stem from at least two mechanisms. First, partners may perceive actual changes in their significant other’s typical pattern of thoughts, feelings, and behaviors. Second, regardless if actual personality change has occurred or not, spouses may develop shared perceptions about the ways in which their personality changed and may consequently rate their partner’s personality traits in a way consistent with their significant others’ self-report, as there is evidence to suggest that partner perceptions are not exclusively based on observable behavior alone (Leising et al., 2014). In other words, partners may communicate the ways in which they are changing in various ways. For example, spouses may not only express personality-relevant behavior but may also verbalize the ways in which they see themselves to their romantic partner.

Relatively few existing studies have looked at self–other agreement in adult personality development (Oltmanns et al., 2019; Watson & Humrichouse, 2006). Our study contributes to this growing body of work showing that romantic partners are able to observe changes in their significant others and revise their perceptions of their partner’s personality. However, further studies addressing the underlying mechanisms by which personality traits change would be useful in understanding the sources of personality change and allow researchers to identify knowledgeable raters who are suited to report on such changes.

Together, the cross-sectional and longitudinal results of self–other personality agreement indicate that romantic partners can report on their significant others’ personalities with high agreement, both at one particular point in time and across time. Yet, agreement was less than perfect for all traits, alluding to the possibility that different reporting methods may yield complementary information. Like past studies, we found significant differences in self–other agreement in personality change across couples and analyzed a set of potentially relevant moderators. Contrary to what we predicted, parenthood status, relationship length, cohabitation, and gender did not appear to influence the agreement of self- and partner-reports over time. This finding was surprising

given the extensive evidence for the moderating role of interpersonal factors (such as duration and type of acquaintanceship) in cross-sectional self–other agreement (Connolly et al., 2007; Watson et al., 2000). Notably our sample size was—by contemporary standards—moderate, and our relationship length variable was, on average, well beyond the window evidenced to exhibit differences in self–other agreement. Most important, we had restricted variance in personality change. As such, the apparent lack of moderator results does not necessarily suggest that there are no effects but rather that the effects may have been too small to detect in the current study. However, as noted above and by previous research, it is possible that self–other agreement stabilizes at a similar level across couples after the beginning months of a relationship because judging the other’s personality is a relatively “simple task” (Allik et al., 2016; Rogers & Biesanz, 2019).

### *Limitations*



Our multi-wave study of self- and partner-reports has several strengths that allowed us to address new and hitherto unanswered questions about personality trait development. However, there were also a number of limitations. First, as mentioned above, the sample size may have limited our ability to detect small effects. Future studies with larger samples should investigate self–other agreement using local structural equation modeling (LSEM; Hildebrandt et al., 2016), which allows for the inclusion and analysis of continuous moderators that may be associated with varying levels of self–other agreement. Second, the relatively short research period of only 1.5 years restricted the degree of change in personality that could occur and be reported on. Future research on self–other agreement across longer intervals would be needed to track longer-term changes in personality traits. Third, we assessed self–other agreement in romantic couples but not in any other types of dyads. As such, we were not able to compare levels of agreement across different types of dyads such as friends or colleagues. Future studies should investigate self–other agreement in change in a variety of dyadic relationships to gain a better understanding of when and how self-reported changes in personality traits can be observed by relevant others. Additionally, it would be useful to include meta-perspective reports—people reporting on how they believe others view them (Vazire & Carlson, 2010). This would allow researchers to assess the level of awareness individuals have of the discrepancy between how they see themselves and how others see them. Thus, perhaps gaps between self- and informant-reports are gaps that people are already cognizant of, instead of representing true disagreement. Finally, using observational or experimental methods could be another way to

address personality ratings without the concern of rater-specific biases.

## Conclusion

In the present research, we examined the convergence between self- and partner-reported personality traits across 1.5 years in a sample of Dutch couples. Overall, partner-reports of personality traits yielded similar rank-order stabilities and—with two exceptions—similar mean-level personality trait trajectories. Moreover, individuals agreed to some extent with their partners on changes in their partner's levels of emotional stability and agreeableness. Together, these findings validate existing self-report research on personality development but also highlight the importance of using multiple methods of assessments to address method-specific biases. Future research is needed to assess self–other agreement in personality change across other kinds of dyads to advance our understanding of the information we can obtain from different types of raters.

## Data accessibility statement

  This article earned Open Data, Open Materials and Preregistered badges through Open Practices Disclosure from the Center for Open Science. The data and materials are permanently and openly accessible at <https://osf.io/z86ku/>

## Acknowledgements

We thank Chris Hopwood and Rick Robins for their edits on an earlier version of this manuscript.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

Data were obtained from a study funded by a grant to Wiebke Bleidorn from the Netherlands Organization for Scientific Research (NWO, VENI project # 451-13-025).

## ORCID iD

Wiebke Bleidorn  <https://orcid.org/0000-0003-3795-8143>

## Notes

1. Branje et al. (2007) studied self- and other-reported personality development in 285 Dutch families across 3 years, thus incorporating both adolescents and adults.
2. Sample sizes per assessment wave, parenthood status, and gender are shown in Table S1 in the online supplementary material.
3. Dyadic correlations for self- and partner-reported personality traits at Wave 1 are presented in Table S5 in the online supplementary material.

4. We initially attempted to also analyze relationship quality as a moderator variable. However, the skewed distribution and restricted range of this variable precluded meaningful moderator tests. We report these analyses and the results in the online supplementary material.

## Supplemental material

Supplemental material for this article is available online.

## References

- Allemand, M., Zimpritch, D., & Hertzog, C. (2007). Cross-sectional age differences and longitudinal age changes of personality in middle adulthood and old age. *Journal of Personality, 75*(2), 323–358.
- Allik, J., de Vries, R. E., & Realo, A. (2016). Why are moderators of self-other agreement difficult to establish? *Journal of Research in Personality, 63*, 72–83. <http://dx.doi.org/10.1016/j.jrp.2016.05.013>
- Beer, A., & Watson, D. (2008a). Asymmetry in judgments of personality: Others are less differentiated than the self. *Journal of Personality, 76*(3), 535–560. <https://doi.org/10.1111/j.1467-6494.2008.00495.x>
- Beer, A., & Watson, D. (2008b). Personality judgment at zero acquaintance: Agreement, assumed similarity, and implicit simplicity. *Journal of Personality Assessment, 90*(3), 250–260. <https://doi.org/10.1080/00223890701884970>
- 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*(6), 723–737. <https://doi.org/10.1177/1073191113500521>
- Biesanz, J. C., & West, S. G. (2004). Towards understanding assessments of the Big Five: Multitrait-multimethod analyses of convergent and discriminant validity across measurement occasion and type of observer. *Journal of Personality, 72*(4), 845–876. <https://doi.org/10.1111/j.0022-3506.2004.00282.x>
- Biesanz, J. C., West, S. G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintance and consensus and self–other agreement in judgments of personality. *Journal of Personality and Social Psychology, 92*(1), 119–135. <https://doi.org/10.1037/0022-3514.92.1.119>
- Bleidorn, W. (2012). Hitting the road to adulthood: Short-term personality development during a major life transition. *Personality and Social Psychology Bulletin, 38*, 1594–1608.
- Bleidorn, W. (2015). What accounts for personality maturation in early adulthood? *Current Directions in Psychological Science, 24*, 245–252.
- Bleidorn, W., Klimstra, T.A., Denissen, J.J.A., Rentfrow, P.J., Potter, J., & Gosling, S.D. (2013). Personality maturation around the world: A cross-cultural examination of Social Investment Theory. *Psychological Science, 24* (12), 2530–2540.
- Bleidorn, W., Hill, P., Back, M., Denissen, J. J. A., Hennecke, M., Hopwood, C. J., Jokela, M., Kandler, C., Lucas, R. E., Luhmann M., Orth, U., Wagner, J., Wrzus, C., Zimmermann, J., & Roberts, B. (2019, June 4). The Policy Relevance of Personality Traits. Advance online publication. <https://doi.org/10.31234/osf.io/a9rbn>

- Bleidorn, W., & Hopwood, C. J. (2019). Stability and change in personality traits over the lifespan. In D. P. McAdams, R. Shiner, & J. Tackett (eds.), *Handbook of personality development* (Vol. 1, pp. 237–252). Guilford.
- Bleidorn, W., Hopwood, C. J., Back, M.D., Denissen J.J.A., Hennecke, M., Jokela, M., Kandler, C., Lucas, R. E., Luhmann M., Orth, U., Roberts, B. W., Wagner, J., Wrzus, C., & Zimmermann, J. (2020). Longitudinal Experience-Wide Association Studies (LEWAS) - A framework for studying personality change. *European Journal of Personality*, *34*, 285–300.
- Bornstein, R.F. (2009). Heisenberg, Kandinsky, and the heteromethod convergence problem: Lessons from within and beyond psychology. *Journal of Personality Assessment*, *91*(1), 1–8.
- Branje, S. J. T., Van Lieshout, C. F. M., & Gerris, J. R. M. (2007). Big five personality development in adolescence and adulthood. *European Journal of Personality*, *21*(1), 45–62. <https://doi.org/10.1002/per.596>
- Chan, M., Rogers, K. H., Parisotto, K. L., & Biesanz, J. C. (2011). Forming first impressions: The role of gender and normative accuracy in personality perception. *Journal of Research in Personality*, *45*(1), 117–120. <https://doi.org/10.1016/j.jrp.2010.11.001>
- Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment*, *15*(1), 110–117. <https://doi.org/10.1111/j.1468-2389.2007.00371.x>
- Dobewall, H., & Aavik, T. (2016). Rank-order consistency and profile stability of self- and informant-reports of personal values in comparison to personality traits. *Journal of Individual Differences*, *37*(1), 40–48. <http://dx.doi.org/10.1027/1614-0001/a000186>
- Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). Quantitative methodology series. An introduction to latent variable growth curve modeling: Concepts, issues, and applications (2nd ed.). Lawrence Erlbaum Associates Publishers.
- Fletcher, G. J. O., & Kerr, P. S. G. (2010). Through the eyes of love: Reality and illusion in intimate relationships. *Psychological Bulletin*, *136*(4), 627–658. <http://dx.doi.org/10.1037/a0019792>
- Funder, D. C. (1995). On the Accuracy of Personality Judgment: A realistic approach. *Psychological Review*, *102*(4), 652–670. <https://doi.org/10.1037/0033-295X.102.4.652>
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, *21*(3), 177–182. <https://doi.org/10.1177/0963721412445309>
- 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*(4), 656–672. <http://dx.doi.org/10.1037/0022-3514.69.4.656>
- Gallrein, A. M. B., Carlson, E. N., Holstein, M., & Leising, D. (2013). You spy with your little eye: People are “blind” to some of the ways in which they are consensually seen by others. *Journal of Research in Personality*, *47*(5), 464–471. <https://doi.org/10.1016/j.jrp.2013.04.001>
- Göllner, R., Roberts, B. W., Damian, R. I., Lüdtke, O., Jonkmann, K., & Trautwein, U. (2017). Whose “storm and stress” is it? Parent and child reports of personality development in the transition to early adolescence. *Journal of Personality*, *85*(3), 376–387. <https://doi.org/10.1111/jopy.12246>
- Hayes, A. F., & Dunning, D. (1997). Construal processes and trait ambiguity: Implications for self-peer agreement in personality judgment. *Journal of Personality and Social Psychology*, *72*(3), 664–677. <https://doi.org/10.1037/0022-3514.72.3.664>
- Hertzog, C., Lindenberger, U., Ghisletta, P., & Von Oertzen, T. (2006). On the power of multivariate latent growth curve models to detect correlated change. *Psychological Methods*, *11*(3), 244–252. <https://doi.org/10.1037/1082-989X.11.3.244>
- Hildebrandt, A., Lüdtke, O., Robitzsch, A., Sommer, C., & Wilhelm, O. (2016). Exploring factor model parameters across continuous variables with local structural equation models. *Multivariate Behavioral Research*, *51*(2–3), 257–258. <https://doi.org/10.1080/00273171.2016.1142856>
- Hofmann, W., Gschwendner, T., & Schmitt, M. (2009). The road to the unconscious self not taken: Discrepancies between self- and observer-inferences about implicit dispositions from nonverbal behavioural cues. *European Journal of Personality*, *23*(4), 343–366. <https://doi.org/10.1002/per.722>
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453. <http://dx.doi.org/10.1037/1082989X.3.4.424>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. <http://dx.doi.org/10.1080/10705519909540118>
- Human, L. J., Carlson, E. N., Geukes, K., Nestler, S., & Back, M. D. (2020). Do accurate personality impressions benefit early relationship development? The bidirectional associations between accuracy and liking. *Journal of Personality and Social Psychology*, *118*(1), 199. <http://dx.doi.org/10.1037/pspp0000214>
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory – versions 4a and 5*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). The Guilford Press.
- 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. <http://dx.doi.org/10.1111/j.1467-6494.1993.tb00781.x>
- Kandler, C., Bleidorn, W., Riemann, R., Spinath, F. M., Thiel, W., & Angleitner, A. (2010). Sources of cumulative continuity in personality: A longitudinal multiple-rater twin study. *Journal of Personality and Social Psychology*, *98*(6), 995–1008. <https://doi.org/10.1037/a0019558>
- Kandler, C., Kornadt, A.E., Hagemeyer, B., & Neyer, F.J. (2015). Patterns and sources of personality development in old age. *Journal of Personality and Social Psychology*, *109*(1), 175–191.



- Kashy, D. A., & Snyder, D. K. (1995). Measurement and data analytic issues in couples research. *Psychological Assessment*, 7, 338–348.
- Kenny, D. A. (1995). The effect of nonindependence on significance testing in dyadic research. *Personal Relationships*, 2, 67–75.
- Kubinger, K. D., & Litzenberger, M. (2006). On the shortcomings of a mono-methodical approach to personality assessment. *Individual Differences Research*, 4(3), 139–158.
- Kurtz, J. E., & Sherker, J. L. (2003, February 1). Relationship quality, trait similarity, and self-other agreement on personality ratings in college roommates. *Journal of Personality*, 71(1), 21–48. <https://doi.org/10.1111/1467-6494.t01-1-00005>
- Leising, D., Erbs, J., & Fritz, U. (2010). The letter of recommendation effect in informant ratings of personality. *Journal of Personality and Social Psychology*, 98(4), 668–682. <http://dx.doi.org/10.1037/a0018771>
- Leising, D., Gallrein, A.-M. B., & Dufner, M. (2014). Judging the behavior of people we know: Objective assessment, confirmation of preexisting views, or both? *Personality and Social Psychology Bulletin*, 40, 153–163. <https://journals.sagepub.com/doi/10.1177/0146167213507287>
- Luan, Z., Hutteman, R., Denissen, J. J. A., Asendorpf, J. B., & van Aken, M. A. G. (2017). Do you see my growth? Two longitudinal studies on personality development from childhood to young adulthood from multiple perspectives. *Journal of Research in Personality*, 67, 44–60. <https://doi.org/10.1016/j.jrp.2016.03.004>
- McCrae, R. R., Stone, S. V., Fagan, P. J., & Costa, P. T. (1998). Identifying causes of disagreement between self-reports and spouse ratings of personality. *Journal of Personality*, 66(3), 285–313. <https://doi.org/10.1111/1467-6494.00013>
- Mroczek, D. K., & Spiro, A. (2003). Modeling intraindividual change in personality traits: Findings from the normative aging study. *Journals of Gerontology – Series B Psychological Sciences and Social Sciences*, 58(3), P153–P165. <https://doi.org/10.1093/geronb/58.3.P153>
- Nave, C. S., Feeney, M. G., & Furr, R. M. (2018). Behavioral observation in the study of personality and individual differences. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *The SAGE handbook of personality and individual differences: The science of personality and individual differences* (pp. 317–340). <https://doi.org/10.4135/9781526451163.n15>
- Neyer, F.J., & Lehnart, J. (2007). Relationships matter in personality development: Evidence from an 8-year longitudinal study across young adulthood. *Journal of Personality*, 75(3), 535–568 .
- Oltmanns, J. R., Jackson, J. J., & Oltmanns, T. F. (2019). Personality change: Longitudinal self-other agreement and convergence with retrospective-reports. *Journal of Personality and Social Psychology*. Advance online publication. <https://doi.org/10.1037/pspp0000238>
- Pusch, S., Mund, M., Hagemeyer, B., & Finn, C. (2019). Personality development in emerging and young adulthood: A study of age differences. *European Journal of Personality*, 33, 245–263. <https://doi.org/10.1002/per.2181>
- R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rau, R., Carlson, E. N., Back, M. D., Barranti, M., Gebauer, J. E., Human, L. J., Leising, D., & Nestler, S. (2019). What is the structure of perceiver effects? On the importance of global positivity and trait-specificity across personality domains and judgment contexts. *Journal of Personality and Social Psychology*. Advance online publication. <https://doi.org/10.1037/pspp0000278>
- Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126(1), 3–25. <https://doi.org/10.1037/0033-2909.126.1.3>
- Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science*, 17(1), 31–35. <https://doi.org/10.1111/j.1467-8721.2008.00543.x>
- Roberts, B.W., & Nickel, L.B. (in press). Personality development across the life course: A neo-socioanalytic perspective. In O.P. John & R.W. Robins (Eds.), *Handbook of Personality Theory and Research*. Guilford Press.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, 132(1), 1–25. <https://doi.org/10.1037/0033-2909.132.1.1>
- Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of Personality: Theory and Research* (pp. 375–398). The Guilford Press.
- Robins, R. W., Fraley, R. C., Roberts, B. W., & Trzesniewski, K. H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality*, 69(4), 617–640. <https://doi.org/10.1111/1467-6494.694157>
- Rogers, K. H., & Biesanz, J. C. (2019). Reassessing the good judge of personality. *Journal of Personality and Social Psychology*, 117(1), 186–200. <https://doi.org/10.1037/pspp0000197>
- Rohrer, J. M., Egloff, B., Kosinski, M., Stillwell, D., & Schmukle, S. C. (2018). In your eyes only? Discrepancies and agreement between self- and other-reports of personality from age 14 to 29. *Journal of Personality and Social Psychology*, 115(2), 304–320. <http://dx.doi.org/10.1037/pspp0000142>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36.
- Schwaba, T., & Bleidorn, W. (2018). Individual differences in personality change across the adult lifespan. *Journal of Personality*, 86(3), 450–464. <https://osf.io/h7apv/>
- 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(2), 330–348. <https://doi.org/10.1037/a0021717>
- Specht, J., Bleidorn, W., Denissen, J. J., Hennecke, M., Hutteman, R., Kandler, C., Luhmann, M., Orth, U., Reitz, A. K., & Zimmermann, J. (2014). What drives adult personality development? A comparison of theoretical perspectives and empirical evidence. *European Journal of Personality*, 28, 216–230. <https://onlinelibrary.wiley.com/doi/10.1002/per.1966>



- Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology, 101*(4), 862–882. <http://dx.doi.org/10.1037/a0024950>
- Tracy, J. L., Robins, R. W., & Sherman, J. W. (2009a). The practice of psychological science: Searching for Cronbach's two streams in social-personality psychology. *Journal of Personality and Social Psychology, 96*, 1206–1225.
- van Scheppingen, M. A., Denissen, J. J. A., & Bleidorn, W. (2018). Stability and change in self-control during the transition to parenthood. *European Journal of Personality, 32*, 690–704. <https://doi.org/10.1002/per.2172>.
- Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality, 40*(5), 472–481. <http://dx.doi.org/10.1016/j.jrp.2005.03.003>
- Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology, 98*(2), 281–300. <https://doi.org/10.1037/a0017908>
- Vazire, S., & Carlson, E. N. (2010). Self-knowledge of personality: Do people know themselves? *Social and Personality Psychology Compass, 4*(8), 605–620. <https://doi.org/10.1111/j.1751-9004.2010.00280.x>
- Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: The role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology, 78*(3), 546–558. <https://doi.org/10.1037/0022-3514.78.3.546>
- Watson, D., & Humrichouse, J. (2006). Personality development in emerging adulthood: Integrating evidence from self-ratings and spouse ratings. *Journal of Personality and Social Psychology, 91*(5), 959–974. <https://doi.org/10.1037/0022-3514.91.5.959>
- 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.
- Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116*(2), 363–381. <http://dx.doi.org/10.1037/0033-2909.116.2.363>
- Wilt, J., & Revelle, W. (2015). Affect, behaviour, cognition and desire in the Big Five: An analysis of item content and structure. *European Journal of Personality, 29*(4), 478–497. <https://doi.org/10.1002/per.2002>
- Wortman, J., Lucas, R. E., & Donnellan, M. B. (2012). Stability and change in the Big Five personality domains: Evidence from a longitudinal study of Australians. *Psychology and Aging, 27*(4), 867–874. <http://dx.doi.org/10.1037/a0029322>
- Zimmermann, J., Schindler, S., Klaus, G., & Leising, D. (2018). The effect of dislike on accuracy and bias in person perception. *Social Psychological and Personality Science, 9*(1), 80–88. <https://doi.org/10.1177/1948550617703167>