Published in final edited form as:

Soc Psychol Personal Sci. 2017 January ; 8(1): 29–35. doi:10.1177/1948550616662026.

Compensatory Conscientiousness Redux: A Direct Replication of Roberts, Smith, Jackson, and Edmonds (2009)

Lauren B. Nickel,

University of Illinois at Urbana-Champaign

James Iveniuk, and

University of Toronto

Brent W. Roberts

University of Illinois at Urbana-Champaign

Abstract

The present study attempted to closely replicate Roberts, Smith, Jackson, and Edmonds (2009) who found, in part, a compensatory effect such that individuals with spouses higher in conscientiousness reported higher self-rated health and fewer physical limitations in the Health and Retirement Study. Using similarly structured data from the National Social Life, Health, and Aging Project ($N = 953$), the current study found results largely consistent with the original study such that partner conscientiousness predicted both self-rated health and physical limitations, and husbands' conscientiousness and neuroticism interacted when predicting wives' self-rated health. A discussion of the usefulness of statistical significance versus effect size in replication follows.

Keywords

conscientiousness; neuroticism; self-rated health; functional limitations; replication

The link between an individual's level of conscientiousness and his or her health has been well-established, with more conscientious individuals exhibiting greater longevity (Friedman et al., 1993), more health-promoting and fewer health-diminishing behaviors (Bogg & Roberts, 2004), and better physical health (Goodwin & Friedman, 2006). However, other important people in an individual's life, such as a spouse or romantic partner, may also have an impact on the individual's health via personality. Traits such as conscientiousness may result in someone not only taking care of themselves better but also their romantic partner. For example, a conscientiousness spouse may be more likely to give reminders to their partner about activities such as taking medications or going to the doctor, which would lead to better health.

Correspondence concerning this article should be addressed to Lauren B. Nickel, Department of Psychology, University of Illinois, 603 E. Daniel St., Champaign, IL 61820. nickel2@illinois.edu.

Author Note

Lauren B. Nickel, Department of Psychology, University of Illinois at Urbana-Champaign; James Iveniuk, Dalla Lana School of Public Health, University of Toronto; Brent W. Roberts, Department of Psychology, University of Illinois at Urbana-Champaign.

A study by Roberts, Smith, Jackson, and Edmonds (2009) used the actor-partner interdependence model as a way to study this relationship, using data from the Health and Retirement Study (HRS), a large nationally representative longitudinal study of Americans 50 years and older and their spouses, who are surveyed every two years to track factors associated with health and retirement. They studied the relationship between an individual's level of conscientiousness, as well as the spouse's level of conscientiousness, and an individual's self-rated health and physical limitations. Findings indicated that a person's health is associated with how conscientiousness his or her spouse is. The husbands of women with higher levels of conscientiousness reported better subjective health and fewer physical limitations; a similar relationship was found between women's subjective health and physical limitations and their husbands' level of conscientiousness. Additionally, there was an interaction between conscientiousness and neuroticism in contributing to a spouse's health; women whose husbands were high in both traits reported better subjective health than other women.

Other studies have used similar models to investigate how an individual's personality traits can influence their spouse's health. For example, in a study of cardiac patients, patients with a more conscientious spouse reported fewer anxiety and depression symptoms (Karademas & Tsousis, 2014). Similarly, individuals with a more optimistic spouse reported better self-rated health and physical functioning (Kim, Chopik, & Smith, 2014). This further indicates the role that the personality of a spouse can play in influencing an individual's health.

While the original idea that a spouse's personality contributes to one's health above and beyond one's own personality has clearly been conceptually replicated, we know of no direct replication of the original effect. Given the increased emphasis on, and the importance of, direct replications (Asendorpf et al., 2013), it would be important to show that this type of pattern could be reproduced in a similar sample using the same measures. Of course, replicating large panel studies is seldom the objective goal of subsequent large panel studies. However, in our experience we have found it to be quite common for research teams to copy methods used in other large panel studies as it is seen as an indicator of using reliable, and proven methods. In particular, this behavior recently resulted in the creation of a second data source, derived entirely independent of the original research team, when the same personality and health measures were used in a survey of the National Social Life, Health, and Aging Project (NSHAP). When we became aware of this second data source, we immediately pre-registered our project (<http://osf.io/bfk24>) and conducted the most direct replication possible of Roberts et al. (2009).

Method

Participants

Our data came from the second wave of the National Social Life, Health and Aging Project (NSHAP) (Waite et al., 2013). NSHAP is a nationally representative study of older adults, designed to collect extensive information on the social lives of older respondents, as well as a broad array of assessments of health. The first wave of NSHAP, collected in 2005–06, comprised 3005 respondents with a response rate of 75.5%. By Wave 2, fielded in 2010–11, 318 respondents became deceased, 115 were in too poor health to be reinterviewed, and an

additional 311 were not reinterviewed for various reasons including refusals. Of those partners who were asked to participate in NSHAP, 84.5% consented and were interviewed, yielding a sample of 955 partners, and thus, 955 marital and cohabitational dyads. Spouses and coresident partners were interviewed using the same protocol as the focal respondents; note that age was not used as a criterion for whether a *partner* would be interviewed, so respondents younger than 57 could be added to the sample. Each participant responded to the survey separately from their partner. There was one same-sex female couple and one same-sex male couple; since this is too few to make inferences about non-heterosexual pairings, we did not include these couples in the analysis below, leaving 953 couples. The Wave 2 response rate was 76.9%, including partners. The average age of participants included in this study was 70.61 ($SD = 7.91$). The average number of years of education for participants in this study was 13.28 ($SD = 3.95$). While the sample size of the NSHAP was smaller than that used in the Roberts et al. (2009) study, which used data from 2,203 couples, it still provided 80% power to detect effects as small as a correlation of .08.

Measures

Personality.—NSHAP contains a 21-item personality battery adapted from the Midlife Development Inventory (MIDI), which has been used in numerous previous surveys (Iveniuk, Laumann, McClintock, Tiedt, & Waite, 2014; Lachman & Weaver, 1997; Zimprich, Allemand, & Lachman, 2012), and is the same personality battery used in Roberts et al. (2009). We obtained the Big Five by creating averaged scales, again, using identical methods to Roberts et al. (2009). Participants were asked to indicate how much each adjective described them on a scale from 1 (*Not at all*) to 4 (*A lot*). The items associated with conscientiousness and neuroticism were used in this study. Conscientiousness was measured using four adjectives: Organized, responsible, hardworking, and thorough. Neuroticism was measured using four items: Moody, worrying, nervous, and calm (reversed).

Health.—Health was assessed using two measures. Self-rated health was obtained by asking “Would you say your health is excellent, very good, good, fair or poor?” To assess functional limitations, respondents were asked how much difficulty they had with the following activities of daily living (ADLs): Walking across a room, walking a block, dressing, bathing, eating, getting in and out of bed, and toileting. Furthermore, respondents were asked how much difficulty they had with the following instrumental activities of daily living (IADLs): Preparing meals, taking medications, managing money, shopping for groceries, performing light housework, and using a telephone. We produced a count of how many of these activities the respondents had any difficulty with; possible responses were *no difficulty*, *some difficulty*, *much difficulty*, or *unable to do*. This count variable could range from 0 to 13. This differs from the functional limitations variable used in Roberts et al. (2009), where the count ranged from 0 to 23. With the exception of walking across a room, all ADLs and IADLS assessed in this study were also assessed in the original study; the original study included additional ADLs and IADLS, such as jogging one mile, climbing up a flight of stairs, and picking up a dime.

Roberts et al. (2009) controlled for age and years of education as variables that might confound the relationship between traits and health. Therefore, we controlled for the age of the focal respondent and the spouse, as well as the total years of education obtained by the focal respondent and the spouse. Additionally, we controlled for the other traits not predicting the outcome variables.

Analyses

We employed Actor-Partner Interdependence Models (APIM; Kenny, 1996) in order to examine the associations between a person's traits and his or her health, as well as his or her spouse's traits on his or her health. APIMs were fit using Seemingly Unrelated Regression Equations (Zellner, 1962, 1963), which simultaneously fit an equation for husbands and an equation for wives, using husbands' and wives' traits in both regressions, as well as estimating the covariances between predictors and the correlation between the error terms in both regression equations. Each case is a couple containing the predictor and outcome variables associated with both the husband and the wife. APIMs were fit with Full Information Maximum Likelihood to assuage problems with missing data (Enders, 2001; Enders & Bandalos, 2001). We standardized all variables within genders before fitting APIMs in order to facilitate comparison of coefficients both within and across models. All analyses were run in Stata version 13 software.

To better understand whether or not this current study replicated the effects of the original study, we compared the effect sizes of the results between the two studies. To do so, we converted the p -values associated with this study's analyses to r , using the effect size calculator found on the Campbell Collaboration's website (<http://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php>) (the Campbell Collaboration is an organization that conducts systematic reviews). The 95% confidence intervals were calculated using a sample size of 953. For comparison, we took the correlations reported in the published manuscript of the original study. To be consistent and help with interpretations, we report only the correlational effect sizes in the results.¹

Results

Descriptive statistics and correlations

Table 1 contains the overall descriptive statistics and zero-order correlations for all variables included in the models. Table 2 contain the descriptive statistics and zero-order correlations for husbands and wives separately.

Does conscientiousness predict better health?

The effect sizes and associated 95% confidence intervals from both the original study and the current study can be seen in Table 3 (for self-rated health) and Table 4 (for functional limitations). Not all correlations were included in the original study's manuscript; when they were available, we calculated whether or not the difference between the correlations were statistically significant using the calculator at <http://vassarstats.net/rdiff.html>. Each line in

¹Beta weights and exact p -values can be found in the OSF repository at <http://osf.io/bfk24>.

the tables represents a separate model. Similar to the original study, actor conscientiousness was positively associated with self-rated physical health in both husbands ($b = 0.32$, $SE = 0.07$, $r_p = .15$, 95% CI [.09, .22]), and wives ($b = 0.32$, $SE = 0.08$, $r_p = .15$, 95% CI [.08, .21]). Actor conscientiousness was negatively associated with functional limitations in both husbands ($b = -0.72$, $SE = 0.19$, $r_p = -.14$, 95% CI [-.20, -.07]), and wives, ($b = -1.14$, $SE = 0.18$, $r_p = -.22$, 95% CI [-.29, -.15]). These main effects of actor conscientiousness on self-reported health and functional limitations were highly similar to those found in the original study. While all of these results replicated using statistical significance as a guide, the effect sizes for the current study were considerably smaller in three out of four cases.

Does partner conscientiousness predict health?

Roberts et al. (2009) also found that the conscientiousness of one's spouse was associated with one's own better ratings of subjective health, which suggests that even if the actor is not highly conscientious, he or she may reap the benefits of conscientiousness if he or she is married to a conscientious person. In this replication study we found that the conscientiousness of husbands mattered for wives' ratings of health. For wives, husbands' conscientiousness was associated with better physical health, ($b = 0.16$, $SE = 0.07$, $r_p = .08$, 95% CI [.01, .15]). However, unlike in the original study we did not find that wives' conscientiousness predicted husbands' subjective ratings of health ($b = -0.02$, $SE = 0.08$, $r_p = .01$, 95% CI [-.06, .08]). The findings for functional limitations more closely matched the original study. As in the original study, wives' conscientiousness did not predict husbands' functional limitations ($b = -0.04$, $SE = 0.19$, $r_p = -.01$, 95% CI [-.08, .06]), but the effects were indistinguishable across studies as the effect sizes fell within their respective confidence intervals (i.e., $r = -.04$ and $-.01$). The relation of husbands' conscientiousness to wives' functional limitations (here, $b = -0.53$, $SE = 0.18$, $r_p = -.10$, 95% [-.17, -.04]) was quite similar across both studies ($r = -.08$ and $-.10$), and both were statistically significant. Therefore, three out of four partner effects replicated.

Is there a synergistic effect of conscientiousness on health?

We also tested whether or not there were interactions between the levels of conscientiousness of husbands and wives, producing better health. In the original study, Roberts et al. (2009) found no significant interactions between actor and partner conscientiousness. In contrast, we found one statistically significant effect in this study: If both members of a couple were highly conscientious, then wives (but not husbands) reported fewer functional limitations ($b = 0.59$, $SE = 0.27$, $r_p = .08$, 95% CI [.01, .15]). This pattern would indicate that, if anything, the current study finds more support for synergistic effects than the original study.

How does neuroticism shape health outcomes?

Neuroticism is also consequential for health and health processes (Lahey, 2009). In the original study, actor neuroticism was negatively related to subjective health ratings and positively related to functional limitations for both husbands and wives. The findings from the current study replicated these effects quite closely. Actor neuroticism was associated with lower ratings of physical health for both husbands ($b = -0.23$, $SE = 0.06$, $r_p = -.14$, 95% CI [-.21, -.07]), and wives ($b = -0.32$, $SE = 0.06$, $r_p = -.19$, 95% CI [-.26, -.12]).

Furthermore, actor neuroticism was positively associated with functional limitations for both husbands ($b = 0.75$, $SE = 0.14$, $r_p = .18$, 95% CI [.12, .25]) and wives ($b = 0.51$, $SE = 0.14$, $r_p = .13$, 95% CI [.06, .20]). Partner neuroticism was not associated with physical health or functional limitations. In terms of statistical significance, all of the effects were replicated. However, like the effects for conscientiousness, most of the effects of neuroticism were smaller in the replication study according to the 95% confidence intervals.

Is there a synergistic effect of conscientiousness and neuroticism?

Finally, we tested whether or not health was predicted by the interactions between conscientiousness and neuroticism. Roberts et al. (2009) found that the interaction of high actor conscientiousness and neuroticism was associated with fewer functional limitations for both husbands and wives. In this study, there was no significant interaction between conscientiousness and neuroticism predicting either physical health or functional limitations, but the effect sizes across both studies were indistinguishable from one another. We would attribute the lack of statistical significance in this case to low power in the replication study.

Roberts et al. (2009) also found that wives reported greater subjective health when their husbands were high in both neuroticism and conscientiousness. We found the same result here; wives with husbands high in both neuroticism and conscientiousness reported better physical health ($b = 0.18$, $SE = 0.08$, $r_p = .08$, 95% CI [.01, .15]).²

Discussion

In the current study, we took advantage of a methodological coincidence: A large national survey (NSHAP) duplicated the basic methods used in prior research using another large national sample (HRS), enabling us to conduct as direct of a replication as possible. Similar to the original study, we found that being married to someone who was more conscientious was associated with rating one's own health as better across several indices. These results indicate that the level of conscientiousness of important others (in this case, the spouse) in a individual's life matters as well, not just the individual's level of conscientiousness. When looking at health, an important factor to consider is the spouse. Presumably, the reason a spouse's personality predicts their partner's health would be that the spouse's personality helps to create an environment that leads to greater health. One could speculate that having a more conscientious partner would lead to better health behaviors, such as eating better and exercising, in part, because the spouse participates in those activities too.

We also found evidence for synergistic effects between conscientiousness and neuroticism, especially for wives. While a spouse's level of neuroticism itself did not have any notable effect on health, the combination of high conscientiousness and high neuroticism in husbands benefitted their wives. While high neuroticism may predict worse health when taken by itself, being married to a man who possesses the combination of the two traits was protective of partner's health. Presumably, the combination of high conscientious and high

²Similar models were also run for the other three Big Five traits: openness, agreeableness, and extraversion. For husbands, there were actor effects for both physical health and functional limitations for all three traits, but no partner effects. For wives, there was an actor effect of extraversion on both physical health and functional limitations, but no partner effects of the other three traits.

neuroticism results in behaviors that are more cautious or more vigilant, resulting in increased attention to their partners' health. The increased concern or worry for one's spouse would be one potential reason for their improved health.

If one takes a broad view of replication, we found highly similar findings to Roberts et al. (2009) in the current study. Of course, the devil is in the details when it comes to replication—the details in this case have to do with the index used to infer that a study or a respective result has replicated. For example, in the recent Reproducibility Report (Open Science Collaboration, 2015), the authors used multiple indices, including statistical significance, effect sizes, and subjective judgments, to infer whether a result replicated.

We argue for two reasons that statistical significance should be the primary criterion used to judge whether the results of Roberts et al. (2009), in particular, were replicated or not. First, the original study was published well prior to the reproducibility crisis in psychology. Though the field should have been well aware of the need to use effect sizes in addition to statistical significance to evaluate an effect (e.g., Wilkinson & TFISI, 1999), papers were being judged primarily on statistical significance in 2009. Second, the reviews, which the current readers are not privy to because the practice in most journals is to keep those private, while predominantly positive, were influenced quite strongly by the statistically significant interaction between husbands' conscientiousness and neuroticism on wives' ratings of health. This finding, at the time, was surprising and noteworthy, which was also a strong criterion for publication in *Psychological Science*. Of course, the effect size for this interaction in its original form was quite small, but that did not influence the opinions of the reviewers.

Using statistical significance as a criterion in the current study revealed a relatively strong pattern of replication, at least for the focal effects. Ten out of fourteen of the originally statistically significant effects were statistically significant in this replication study despite having somewhat lower power to detect small effects. Moreover, the focal effects of partner conscientiousness and the interaction between husbands' conscientiousness and neuroticism were all replicated with the exception of one finding. If one included the replication of non-significant effects, the pattern looks even stronger as 12 of 14 non-significant effects were not significant in the current study.

Of course, in an ideal science, we would move away from solely relying on statistical significance, and in most cases, move to identifying the reliable effect size associated with the study idea (Asendorpf et al., 2013). Possibly at some point in the future, psychology will arrive at a time when this type of information has accumulated for most effects because direct replications have become more commonplace. In the meantime, we can compare the effect sizes of past and current research as an adjunct to statistical significance. Using effect sizes as the criterion leads to a similar conclusion, albeit with a bit more complexity. Using the confidence intervals from the current study, 9 of the 14 effect sizes were similar in magnitude, with two interesting caveats. First, many of the effects that failed to replicate according to the confidence interval criterion were actually statistically significant but smaller in magnitude than the original effects. Second, several of the effect sizes that were not statistically significant were indistinguishable in magnitude from the original study.

Of course, in a utopian vision of science, neither the resulting statistical significance nor the effect size would be used to evaluate the publishability of a project (e.g., the journal *Comprehensive Results in Social Psychology*). Rather, as is seen in the move to publish replication reports across multiple journals (e.g., Simons, 2014), papers and projects could be evaluated simply on (1) the level of interest in the idea being tested and (2) the quality of the methods used to test the idea before learning what the result were. It would be optimistic to think that psychology could move to that ideal in a short period of time. Yet this is one reason why we have endeavored to conduct this direct replication as it shows the reality of replication, which will often add complexity in the short run, but hopefully over the long run, it will lead to a strong, reliable body of evidence for and against different ideas.

This study contained many of the same limitations as the original study. First, as discussed in Roberts et al. (2009), this was a cross-sectional study that did not take into account many of the processes that happen over time that may contribute to changes in personality and health, as well as the influences of spouses on one another. Second, this study also used older adults as participants. This is notable because people become more conscientious as they age (Roberts, Walton, & Viechtbauer, 2006), which could result in a different impact of a spouse's personality as the couple ages. Additionally, older couples, because of the time they have spent together, may differ in important ways from their younger or middle-aged counterparts in terms of how an individual's personality affects their spouse's health. Additionally, using different measures may have yielded different results. The measure used in the original study and the current study was short; a longer measure may have provided a different focus. A measure that assesses the facets of conscientiousness may have provided more information about what is the important facet (or facets) that influences a spouse's health—for example, is it more important to be organized, or it is more important to be industrious? The measures are also all self-report measures. When considering the impact of one person's personality on another's health, other-reports of personality may be helpful in knowing how that person's personality is actually perceived. More objective measures of health, such as blood pressure, may also provide important information about the effect of another's personality.

Conclusion

In this direct replication of Roberts et al. (2009), we found very similar results: (1) Conscientiousness and emotional stability predicted better ratings of health; (2) partner conscientiousness predicted better self-ratings of health in most cases; and (3) the most notable result of the original study, husbands' conscientiousness and neuroticism interacted to predict better ratings of health in wives was also replicated. While the study did not conform to some ideal replication recommendations (e.g., having 2.5 times the number of subjects in the original study; Simonsohn, 2015), the unique structure of the data provided an ideal opportunity to conduct as direct of a replication possible of an effect that would presumably be quite difficult to test otherwise.

References

- Asendorpf JB, Conner M, De Fruyt F, De Houwer J, Denissen JJA, Fiedler K, ..., Wicherts JM (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, 27, 108–119. doi: 10.1002/per.1919
- Bogg T, & Roberts BW (2004). Conscientiousness and health behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin*, 130, 887–919. doi: 10.1037/0033-2909.130.6.887 [PubMed: 15535742]
- Enders CK (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling*, 8, 128–141. doi: 10.1207/S15328007SEM0801_7
- Enders CK, & Bandalos DL (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, 8, 430–457. doi: 10.1207/S15328007SEM0803_5
- Friedman HS (2000). Long-term relations of personality and health: Dynamisms, mechanisms, tropisms. *Journal of Personality*, 68, 1089–1107. doi: 10.1111/1467-6494.00127 [PubMed: 11130733]
- Friedman HS, Tucker JS, Tomlinson-Keasey C, Schwartz JE, Wingard DL, & Criqui MH (1993). Does childhood personality predict longevity? *Journal of Personality and Social Psychology*, 65, 176–185. doi: 10.1037/0022-3514.65.1.176 [PubMed: 8355139]
- Goodwin RD, & Friedman HS Health status and the five-factor personality traits in a nationally representative sample. *Journal of Health Psychology*, 11, 643–654. doi: 10.1177/1359105306066610
- Iveniuk J, Laumann EO, McClintock MK, Tiedt A, & Waite LJ (2014). Personality measures in the National Social Life, Health and Aging Project. *Journals of Gerontology Series B: Psychological and Social Sciences*, 69, S117–S124. doi: 10.1093/geronb/gbu073
- Karademas EC, & Tsaousis I (2014). The relationship of patient and spouse personality to cardiac patients' health: Two observational studies of mediation and moderation. *Annals of Behavioral Medicine*, 47, 79–91. doi: 10.1007/s12160-013-9523-5 [PubMed: 23780734]
- Kenny DA (1996). Models of nonindependence in dyadic research. *Journal of Social and Personal Relationships*, 13, 279–294. doi: 10.1177/0265407596132007
- Kim ES, Chopik WJ, & Smith J (2014). Are people healthier if their partners are more optimistic? The dyadic effect of optimism on health among older adults. *Journal of Psychosomatic Research*, 76, 447–453. doi: 10.1016/j.jpsychores.2014.03.104 [PubMed: 24840138]
- Lachman ME, & Weaver SL (1997). The Midlife Development Inventory (MIDI) personality scales: Scale construction and scoring (Technical Report). Brandeis University, Waltham, MA.
- Lahey BB (2009). Public health significance of neuroticism. *American Psychologist*, 64, 241–256. doi: 10.1037/a0015309 [PubMed: 19449983]
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349, aac4716. doi: 10.1126/science.aac4716 [PubMed: 26315443]
- Roberts BW, Kuncel NR, Shiner R, Caspi A, & Goldberg LR (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2, 313–345. doi: 10.1111/j.1745-6916.2007.00047.x [PubMed: 26151971]
- Roberts BW, Smith J, Jackson JJ, & Edmonds G (2009). Compensatory conscientiousness and health in older couples. *Psychological Science*, 20, 553–559. doi: 10.1111/j.1467-9280.2009.02339.x [PubMed: 19476589]
- Roberts BW, Walton KE, & Viechtbauer W (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, 132, 1–25. doi: 10.1037/0033-2909.132.1 [PubMed: 16435954]
- Shanahan MJ, Hill P, Roberts BW, Eccles J, & Friedman HS (2014). Conscientiousness, health and aging: The life course of personality model. *Developmental Psychology*, 50(5), 1407–1425. doi: 10.1037/a0031130 [PubMed: 23244406]
- Simons DJ (2014). The value of direct replication. *Perspectives on Psychological Science*, 9, 76–80. doi: 10.1177/1745691613514755 [PubMed: 26173243]

- Simonsohn U (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, 26, 559–569. doi: 10.1177/0956797614567341 [PubMed: 25800521]
- Waite LJ, Cagney KA, Cornwell B, Dale W, Huang E, Laumann EO, et al. (2013). The National Social Life, Health and Aging Project (NSHAP): Wave 2 and Partner Data Collection.
- Wilkinson L, & APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594–604. doi: 10.1037/0003-066X.54.8.594
- Zellner A (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, 57, 348–368. doi: 10.1080/01621459.1962.10480664
- Zellner A (1963). Estimators for seemingly unrelated regression equations: Some exact finite sample results. *Journal of the American Statistical Association*, 58, 977–992. doi: 10.1080/01621459.1963.10480681
- Zimprich D, Allemand M, & Lachman ME (2012). Factorial structure and age-related psychometrics of the MIDUS personality adjective items across the life span. *Psychological Assessment*, 24, 173–186. doi: 10.1037/a0025265 [PubMed: 21910548]

Table 1.

Descriptive statistics and zero-order correlations for all participants.

Variable	1	2	3	4	5	6	7	8	9
1. Conscientiousness	—								
2. Neuroticism	-.13	—							
3. Openness	.46	-.16	—						
4. Extraversion	.44	-.14	.51	—					
5. Agreeableness	.39	-.07	.33	.55	—				
6. Self-rated health	.22	-.20	.13	.18	.03	—			
7. Functional limitations	-.21	.16	-.08	-.13	-.03	-.43	—		
8. Years of education	.15	-.15	.14	.01	-.01	-.09	-.12	—	
9. Age	-.08	-.05	-.09	-.08	-.05	.24	.17	-.07	—
Mean	2.37	1.24	1.92	2.22	2.44	3.22	1.37	13.28	70.61
SD	0.55	0.61	0.65	0.55	0.54	1.05	2.43	3.95	7.91
Alpha	0.69	0.65	0.75	0.75	0.77	—	—	—	—

Table 2.

Descriptive statistics and zero-order correlations for husbands (below the diagonal) and wives (above the diagonal).

Variable	1	2	3	4	5	6	7	8	9
1. Conscientiousness	—	-.08	.40	.37	.33	.24	-.26	.19	-.05
2. Neuroticism	-.21	—	-.11	-.10	-.09	-.22	.14	-.12	-.06
3. Openness	.52	-.21	—	.43	.31	.17	-.10	.17	-.09
4. Extraversion	.48	-.21	.58	—	.51	.17	-.15	.04	-.05
5. Agreeableness	.39	-.12	.35	.56	—	.08	-.07	-.00	-.04
6. Self-rated health	.22	-.20	.13	.18	.03	—	-.45	.32	-.10
7. Functional limitations	-.18	.20	-.08	-.12	.00	-.43	—	-.13	.16
8. Years of education	.13	-.16	.11	-.00	-.03	.24	-.12	—	-.11
9. Age	-.06	-.06	-.07	-.04	.05	-.09	.19	-.05	—
Mean (husbands)	2.29	1.17	1.90	2.13	2.29	3.18	1.43	13.31	72.36
SD (husbands)	0.57	0.63	0.67	0.56	0.56	1.06	2.42	4.21	7.35
Alpha (husbands)	0.69	0.65	0.77	0.75	0.76	—	—	—	—
Mean (wives)	2.44	1.30	1.94	2.31	2.60	3.26	1.31	13.24	68.84
SD (wives)	0.51	0.59	0.63	0.51	0.46	1.05	2.42	3.67	8.07
Alpha (wives)	0.68	0.66	0.73	0.74	0.73	—	—	—	—

Table 3.

Effect sizes, confidence intervals, and difference between r_p s (where available) for models predicting self-rated health.

	Husbands				Wives			
	Original study		Current study		Original study		Current study	
	r_p	95% CI	r_p	95% CI	r_p	95% CI	r_p	95% CI
Actor C	.20	[.16, .24]	.15	[.09, .22]	.26	[.22, .30]	.15	[.08, .21]
Partner C	.09	[.05, .13]	.01	[-.06, .08]	.07	[.03, .11]	.08	[.01, .15]
Actor N	-.21	[-.25, -.17]	-.14	[-.21, -.07]	-.22	[-.26, -.18]	-.19	[-.26, -.12]
Partner N	n.s.		.07	[-.00, .14]	n.s.		-.03	[-.10, .04]
Actor C $\times$ Partner C	n.s.		.03	[-.04, .10]	n.s.		-.00	[-.07, .07]
Actor C $\times$ Actor N	n.s.		.01	[-.06, .08]	n.s.		-.03	[-.10, .04]
Partner C $\times$ Partner N	n.s.		.03	[-.05, .10]	.05	[.01, .09]	.08	[.01, .15]
								$z = -0.78, p = .44$

Table 4.

Effect sizes, confidence intervals, and difference between r_p s (where available) for models predicting functional limitations.

	Husbands				Diff. between r_p s	Wives				
	Original Study	95% CI	Current Study	95% CI		Original Study	95% CI	Current Study	95% CI	Diff. between r_p s
Actor C	-.20	[-.24, -.16]	-.14	[-.20, -.07]	$z = -1.59, p = .11$	-.23	[-.27, -.19]	-.22	[-.29, -.15]	$z = -0.27, p = .79$
Partner C	-.04	[-.08, .00]	-.01	[-.08, .06]	$z = -0.77, p = .44$	-.08	[-.12, -.04]	-.10	[-.17, -.04]	$z = 0.52, p = .60$
Actor N	.21	[.17, .25]	.18	[.12, .25]	$z = 0.80, p = .42$	.22	[.18, .26]	.13	[.06, .20]	$z = 2.39, p = .02$
Partner N	n.s.		.04	[-.03, .11]		n.s.		.02	[-.09, .05]	
Actor C $\times$ Partner C	n.s.		.02	[-.04, .10]		n.s.		.08	[.01, .15]	
Actor C $\times$ Actor N	-.05	[-.09, -.01]	-.06	[-.13, .01]	$z = 0.26, p = .79$	-.05	[-.09, -.01]	-.04	[-.11, .03]	$z = -0.26, p = .79$
Partner C $\times$ Partner N	n.s.		.05	[-.02, .12]		n.s.		.02	[-.04, .10]	