Iveniuk, J., Laumann, E.O., Waite, L.J., McClintock, M.K., & Tiedt, A. (2014). Personality measures in the national social life, health, and aging project. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *69(8)*, S117–S124, doi:10.1093/geronb/gbu073

Personality Measures in the National Social Life, Health, and Aging Project

James Iveniuk,[1] Edward O. Laumann,[1] Linda J. Waite,[2] Martha K. McClintock,[3] and Andrew Tiedt[4]

[1]Department of Sociology,
[2]Lucy Flower Professor in Urban Sociology, and
[3]Institute for Mind and Biology, and Department of Human Development, University of Chicago, Illinois.
[4]Department of Justice, Bureau of Justice Statistics, Washington, District of Columbia.

Objectives. Provide recommendations for researchers on the use of the Big Five personality battery in the National Social Life, Health, and Aging Project (NSHAP), and ensure that the battery does proxy the Big Five. Also, describe the levels of Big Five traits across gender and age.

Method. We used an Exploratory Structural Equation Model (ESEM) to analyze NHSAP's personality battery, comparing NSHAP with the National Longitudinal Study of Midlife in the United States (MIDUS) and the Health and Retirement Study (HRS).

Results. ESEM revealed a 5-factor structure in the NSHAP battery, but with considerable cross-loadings. When these cross-loadings were not included in the model, model fit notably worsened. Reliabilities of Big Five scales were comparable to the HRS and MIDUS, even though NSHAP's battery is shorter. Women were considerably more Agreeable than men, although this gender gap closed among the oldest in the sample (80 years or older).

Discussion. Researchers will be able to make use of NSHAP's personality battery to examine a range of social, biological, and psychological factors at older ages, in light of individuals' general traits. We recommend models which allow for cross-loadings.

Key words: Big Five—Factor analysis—NSHAP—Personality traits.

THE purpose of this paper is to analyze variation in National Social Life, Health, and Aging Project (NSHAP's) battery of personality items. We will show that the variation in these items is primarily explained by five factors which correspond closely to what have been identified as the Big Five, a commonly used model of personality (John, Naumann, & Soto, 2008; Marsh, Nagengast, & Morin, 2013; Specht, Egloff, & Schmukle, 2011). Following this analysis, we provide recommendations to researchers on how to use this battery in research on aging, traits, health, and social connectedness. The presence of the Big Five in NSHAP provides researchers with a new and unique opportunity to link personality traits and social factors together in terms of how they produce good health at older ages, and this paper provides researchers with methodological guidance for such analyses.

Personality traits describe characteristic differences in individuals' thinking, emotion, and behavior in relation to their environment (Eysenck, 1998 [1947]; Funder, 2001; Roberts, 2009). In the past 10 years, a large and growing body of literature has linked differences in personality traits to differences in physical and mental health (Hampson, 2012; Ozer & Benet-Martínez, 2006; Shannahan, Hill, Roberts, & Eccles, 2012). The effect of Conscientiousness on mortality is comparable to that of socioeconomic status (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007), and there is a strong and positive association between Extraversion and subjective well-being (Ozer & Benet-Martínez, 2006). Overall, research has shown that personality matters across a range of social, psychological, and physical domains (Hampson, 2012; Hampson, Goldberg, Vogt, & Dubanoski, 2007; Shannahan et al., 2012). Furthermore, because traits describe the way that a person interacts with his or her environment, the features of this environment will be consequential for how personality does or does not contribute to good health and well-being (Eysenck 1998 [1947]; Roberts, 2009; Shannahan et al., 2012). Findings that emphasize this perspective, which sees personality and the environment in terms of dynamic relationships, may be enriched by data on the characteristics of individuals' social environments, linked to data on personality traits (Lodi-Smith & Roberts, 2007; Roberts, 2007, 2009; Shannahan et al., 2012).

The present paper describes nationally representative data on personality traits, available from the NSHAP, which may be uniquely valuable to researchers interested in integrated, interdisciplinary models of the aging process because of its rich set of measures on networks, neighborhoods, social support, and marital quality (Cornwell, Laumann, & Schumm, 2008; Cornwell & Waite, 2009; Lindau, Laumann, Levinson, & Waite, 2003). Our aim is

Published by Oxford University Press on behalf of the Gerontological Society of America 2014.
Received November 4, 2013; Accepted May 13, 2014
Decision Editor: Mark Hayward, PhD

to provide researchers with a validation of these measures within the NSHAP sample, since they were modified from previous versions, and also to recommend methods for using these measures in future papers. The first wave of NSHAP produced a number of findings related to older adults' social networks, psychological well-being, and physical health (Cornwell et al., 2008; Cornwell & Waite, 2009, 2012; Lindau et al., 2007). However, the first wave did not include measures of psychological traits. NSHAP Wave 2 includes a personality battery, adapted from the Midlife Development Inventory (MIDI), that proxies the Big Five (Costa & McCrae, 1992a; John et al., 2008; Lachman & Weaver, 1997).

The Big Five

The Big Five traits are: Openness to experience, or seeking variety in types of experiences; Conscientiousness, or being norm-abiding and delaying gratification; Extraversion, or being lively and sociable; Agreeableness or being communally oriented and altruistic; and finally Neuroticism or emotional instability, which entails a vulnerability to negative affect (John et al., 2008). These traits can be abbreviated into the mnemonic acronym "OCEAN."

The Big Five is perhaps the most well-validated taxonomic system for personality, vis-a-vis four important criteria. First, while there may be traits that are beyond or outside of the Big Five, many other traits are nested within the Big Five as subscales or sub-dimensions of the space of possible traits (Costa & McCrae, 1992a; Digman, 1990). Second, the Big Five components can be found across social status, ethnicity, gender, language, or culture, (Costa & McCrae, 1992a; Schmitt, Allik, McCrae, & Benet-Martínez, 2007). Third, Big Five traits have strong interrater reliability, as well as agreement between self-ratings and outside observers (Costa & McCrae, 1992a). Fourth, Big Five traits are highly stable over the life course, especially at older adulthood, meaning it is plausible to assume for most of the NSHAP sample that their traits have been stable for some time before they were interviewed (Cobb-Clark & Schurer, 2012; Roberts & DelVecchio, 2000). Existing work suggests that test–retest correlations of Big Five traits asymptotically approach 1.0 as age increases, and the gap between test and retest decreases (Roberts & DelVecchio, 2000). The Big Five is therefore an integrated taxonomic system for classifying individuals according to personality traits, allowing researchers to proxy factors which are temporally (and perhaps causally) prior to some outcome of interest.

Method

The second wave of NSHAP was collected in 2010–2011, 5 years after the first wave in 2005–2006. The first wave interviewed 3,005 older adults, with a response rate of 75.5%. Seven hundred and forty-four respondents were not reinterviewed at W2, for various reasons including death and poor health. This left 2,261 returning respondents. However, the second wave also instituted a partner module, where 955 spouses or cohabiting romantic partners of respondents were given the same survey instrument as W1 returning respondents. Some respondents who declined to participate in W1, and who were not coded as hostile by the interview team in W1, were also contacted again to see if they would reconsider participating in the survey. Altogether, including partners and new respondents, interviews were completed with 3,377 respondents in W2. The data collection process at both waves involved a 90-min in-person interview, which captured the majority of biological, psychological, and social measures. Following this, the respondent was given a leave-behind questionnaire, to return on their own time, which included the MIDI. The NSHAP MIDI is a list of 21 adjectives that the respondent is asked to use to rate themselves, on a four-point scale, ranging from "not at all" to "a little," "some," and "a lot." In households where both partners participated, both were asked to fill out the leave-behind questionnaire.

It is worth describing how the MIDI compares with other personality batteries that also measure the Big Five, since there are multiple instruments which purport to capture these same traits (Costa & McCrae, 1992b; Gosling, Rentfrow, & Swann, 2003; Lachman & Weaver, 1997). Using personality batteries as part of a omnibus in-home survey creates a problem, because these batteries tend to be quite long (Costa & McCrae, 1992b), and therefore burdensome. Shortening these batteries can also lead to severe problems with reliability and validity (Gosling et al., 2003). The MIDI is relatively short, but according to previous analyses of its factor structure, long enough that it accurately taps the traits of interest, according to numerous previous tests of the measure's validity (Lachman & Weaver, 1997; Zimprich, Allemand, & Lachman, 2012). The developers of the MIDI carried out a pilot study of 574 men and women, and then selected a list of adjectives with high item-to-total correlations and factor loadings for use in the Midlife in the United States (MIDUS) study (Goodwin & Gotlib, 2004). The NSHAP team further modified the MIDI to shorten it further, in order to fit it into an already lengthy leave-behind questionnaire. The differences between the MIDUS MIDI, Health and Retirement Study (HRS) MIDI, and the NSHAP MIDI battery are shown in Table 1. In NSHAP W2, 2,949 individuals returned the leave-behind questionnaire (87.3%), and 2,848 individuals filled out any of the MIDI questions (84.3%). The adjective with the least amount of missing data was "responsible," which 2,754 people answered, and the adjectives with the most missing data were "nervous" and "warm," tied with 2,688 nonmissing values each. After dropping those cases which had missing values on all adjectives (529 respondents), 2,193 respondents had no missing values out of 2,848 (77.0%). Data are publicly available (NSHAP Wave 2: Waite, Linda J., Kathleen Cagney, William Dale, Elbert Huang, Edward O. Laumann, Martha K.

Table 1. Comparison of the MIDI Personality Battery Across Three Nationally Representative Surveys

	MIDUS	HRS	NSHAP
Openness			
Creative	•	•	•
Imaginative	•	•	•
Curious	•	•	•
Adventurous	•	•	•
Intelligent	•	•	
Broad minded	•	•	
Sophisticated	•	•	
Conscientiousness			
Organized	•	•	•
Responsible	•	•	•
Hardworking	•	•	•
Thorough	•	•	•
Careless	•	•	
Extraversion			
Outgoing	•	•	•
Friendly	•	•	•
Lively	•	•	•
Active	•	•	•
Talkative	•	•	•
Agreeableness			
Warm	•	•	•
Caring	•	•	•
Softhearted	•	•	•
Sympathetic	•	•	•
Helpful	•	•	
Neuroticism			
Moody	•	•	•
Worrying	•	•	•
Nervous	•	•	•
Calm	•	•	•

Notes. HRS = Health and Retirement Study; MIDI = Midlife Development Inventory; MIDUS = Midlife in the United States; NSHAP = National Social Life, Health, and Aging Project.
•indicates adjective is present in that survey's version of the MIDI.

McClintock, Colm A. O'Muircheartaigh, L. Phillip Schumm, and Benjamin Cornwell. National Social Life, Health, and Aging Project (NSHAP): Wave 2 and Partner Data Collection. ICPSR34921-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-04-29. doi:10.3886/ICPSR34921.v1.).

Results

Factor Structure of the NSHAP MIDI

We now turn to an analysis of the MIDI factor structure. We carried out an Exploratory Structural Equation Modeling (ESEM) analysis on the MIDI personality battery, in order to ascertain that a Big Five factor structure existed within these adjectives after they had been modified by the NSHAP investigative team. We chose ESEM because other authors have argued that confirmatory factor analyses (CFAs) are not adequate to the complex structure of the Big Five personality traits (Marsh et al., 2010, 2013). All analyses were carried out using MPlus v.6.2 statistical software using mean and variance-adjusted weighted least squares (WLSMV), (Muthén & Muthén, 1998–2011), because WLSMV is appropriate for binary or ordinal outcome variables in an ESEM context (Muthén & Muthén, 1998–2011; Rhemtulla, Brosseau-Liard, & Savalei, 2012). Exploratory factor analyses (EFAs) using Maximum Likelihood produced very similar results, with only five factors having eigenvalues over 1.0. EFA with WLSMV produced identical results, which is expected, since ESEM is EFA in an SEM framework (Asparouhov & Muthén, 2009). WLSMV handles missing data by computing pairwise polychoric correlations between all variables in the data set, rather than restricting the analysis to individuals with complete information on every variable, and this procedure has been shown to be more efficient than listwise deletion (Asparouhov & Muthén, 2010; Flora & Curran, 2004). Factor scores can therefore be estimated even for those who did not answer some of the items, though with less precision. Latent factors predicted manifest, measured adjectives using ordinal probit links—which assumes that each variable reflects an underlying, normally distributed variable—and the factor analysis underwent a varimax rotation. The root mean squared error of approximation (RMSEA) was below 0.06 when there were at least five factors (0.056); the comparative fit index (CFI) and Tucker–Lewis index (TLI) were 0.97 and 0.94 respectively, indicating good model fit. Chi-squared statistics comparing the fitted model to a saturated model were significant up to 10 factors, which is not surprising given the chi-squared test of model fit is sensitive to sample size, and therefore is almost always significant with large samples (Ullman & Bentler, 2003). Although the fit is less than perfect, a five-factor solution fits the data reasonably well, and is commensurate with the expectation that the MIDI does proxy the Big Five. We therefore chose a five-factor solution.

Table 2 shows the varimax-rotated factor loadings. Cells were left blank if factor loadings were not significant at $p < .05$; entries are bolded for the largest factor loading in each row. Note as well that there are two exceptions to the expected factor structure. First, "active," which is supposed to load on to Extraversion, loads more strongly on to Conscientiousness. This may be because at older ages "staying active" is a phrase that is often associated with exercise, and as such, may be more readily understood as an expression of conscientiousness (Mitka, 2001). Second, "warm" seems to load on to Extraversion, although it has a very similar, and almost as large, factor loading on Agreeableness, its expected factor. For the most part, though, the factor structure presents as expected. Splitting the sample by gender also produced very similar results. We therefore hold that the MIDI appropriately captures the Big Five in NSHAP.

Researchers could also carry out a CFA, which constrains factor loadings to zero if the adjective does not load on its theoretically specified latent factor (i.e., the adjective "active" could be constrained to only load on the

Table 2. Varimax-Rotated, Exploratory Structural Equation Model of the MIDI Battery in W2 of NSHAP (n = 2,848)

	F1	F2	F3	F4	F5
Openness					
Creative	**.77**	.17	.16	.11	
Imaginative	**.83**	.15	.15	.20	
Curious	**.44**	.16	.19	.19	.04
Adventurous	**.53**	.24	.40		−.09
Conscientiousness					
Organized	.13	**.63**	.07	.14	
Responsible	.18	**.61**	.14	.41	
Hardworking	.37	**.55**	.19	.15	.05
Thorough	.27	**.61**	.13	.27	−.05
Extraversion					
Outgoing	.21	.06	**.71**	.19	−.06
Friendly	.11	.13	**.68**	.50	−.09
Lively	.28	.32	**.61**	.18	
Active	.31	**.52**	.41		−.07
Talkative	.23	.05	**.58**	.24	.10
Agreeableness					
Warm	.10	.11	**.60**	.57	−.08
Caring	.08	.24	.30	**.69**	
Softhearted	.20	.09	.15	**.73**	.10
Sympathetic	.16	.16	.24	**.73**	.09
Neuroticism					
Moody		−.09	−.10	−.14	**.46**
Worrying			−.06	.11	**.76**
Nervous				.05	**.81**
Calm	.23	.27	.12	.35	**−.41**

Notes. Cells left blank if not significant at $p < .05$. Cells bolded to indicate largest loading for that row.

latent factor for Extraversion, and have no cross-loadings). However, such a model could be a worse fit to the data, and therefore be less effective at capturing the latent factors. To test this, we used an EFA in a CFA framework, which allowed cross-loadings. We then calculated a chi-squared test comparing the fit of two, nested models to each other, one with cross-loadings, and one without. The chi-squared test was significant at $p < .001$, meaning we can confidently reject the null hypothesis that allowing for cross-loadings does not improve model fit. Furthermore, the constrained EFA-in-CFA had notably worse model fit according to the CFI (0.83), the TLI (0.80), and the RMSEA (0.10). We therefore concluded that factor scores from a model which allows for cross-loadings, were a better fit to the data, and more useful to researchers. (Previous studies have also shown that the MIDI factor structure is largely invariant across age groups, recommending it for use among older adults [see Zimprich et al., 2012].)

Comparison of Factor Scores Versus Scales

We compared the factor scores from this analysis to scales constructed from the adjectives; when we say "scores" we mean predicted values of the latent variables produced from the factor analysis in Table 2, while "scales" are the unweighted average of the adjectives that theoretically load onto the latent factors (see Table 1). Alpha reliabilities for these scales were generally good and comparable to alphas from other nationally representative surveys that employ the MIDI. In Table 3, we show the results from calculating the alpha values for personality scales from the second wave of the MIDUS study, as well as the 2010 wave of the HRS. We used the entire sample of each survey's wave to compute these results. We can see that Neuroticism in NSHAP had the lowest alpha value, and was lower than expected given the alphas in MIDUS and HRS, however, it is still within an acceptable range. Note the difference from MIDUS and HRS, suggesting that shortening the scale may have led to this drop in reliability. We also tested the correlations between the more restricted scales, used in NSHAP, and the fuller scales using data from the 2010 wave of the HRS. Looking at Table 1, we can see that the scales for Openness, Conscientiousness, and Agreeableness are shorter in NSHAP than they are in HRS.

Table 3. Comparisons of Alpha Reliabilities for Scales as Found in Wave Two of NSHAP, Wave Two of MIDUS, and the 2010 Wave of the HRS

	NSHAP W2	MIDUS W2	HRS 2010
Openness	.74	.73	.79
Conscientiousness	.69	.68	.68
Extraversion	.75	.76	.75
Agreeableness	.76	.78	.79
Neuroticism	.64	.74	.71

Notes. HRS = Health and Retirement Study; MIDUS = Midlife in the United States; NSHAP = National Social Life, Health, and Aging Project.

We also looked at the intercorrelations within scales, and the correlation between scores and scales; these are shown in Table 4. This table contains two separate correlation matrixes. The first panel shows correlations among the scales, and one result is that the scales are all intercorrelated. It is possible that the varimax-rotated factor scores and the scales do not capture the same latent construct. This could be the case since some adjectives, like "active" and "warm," do not load on to the factor that we would theoretically expect. In order to test this, panel 2 correlates scores with scales. We can see that all correlations along the diagonal—that is, between a score and its scale counterpart—are greater than .90, and in the case of Neuroticism, the correlation is almost perfect. This may be because Neuroticism is the most distinct factor from the other four, in that it has the most minimal cross-loading with other factors. One reason for this might be that the neuroticism scale is composed of three negative items and one positive item, whereas all the other scales are composed of positive items.

Given the high correlation between scales and scores for the same construct, as well as potential problems arising from collinearity among scales, we recommend the use of factor scores. Fitting an SEM with WLSMV, or some other procedure capable of assuaging problems with missing data, could also prove useful when computing factor scores. In addition to the method presented here, a researcher could employ

Table 4. Correlation Matrixes Among Scales, and Correlations Between Factor Scores and Scales in NSHAP W2

Panel 1. Correlation matrix: scales					
	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Openness	1.00				
Conscientiousness	.48***	1.00			
Extraversion	.53***	.46***	1.00		
Agreeableness	.36***	.41***	.55***	1.00	
Neuroticism	−.13***	−.13***	−.15***	−.10***	1.00
Panel 2. Correlation matrix: factor scores with scales					
	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
Factor scores					
Openness	**.95*****	.37***	.38***	.20***	−.09***
Conscientiousness	.24***	**.92*****	.30***	.21***	−.14***
Extraversion	.33***	.21***	**.91*****	.44***	−.10***
Agreeableness	.19***	.30***	.34***	**.94*****	−.10***
Neuroticism	−.05	−.04	−.06	.04*	**.97*****

Notes: Bold indicates main diagonal.
*$p < .05$. ***$p < .001$.

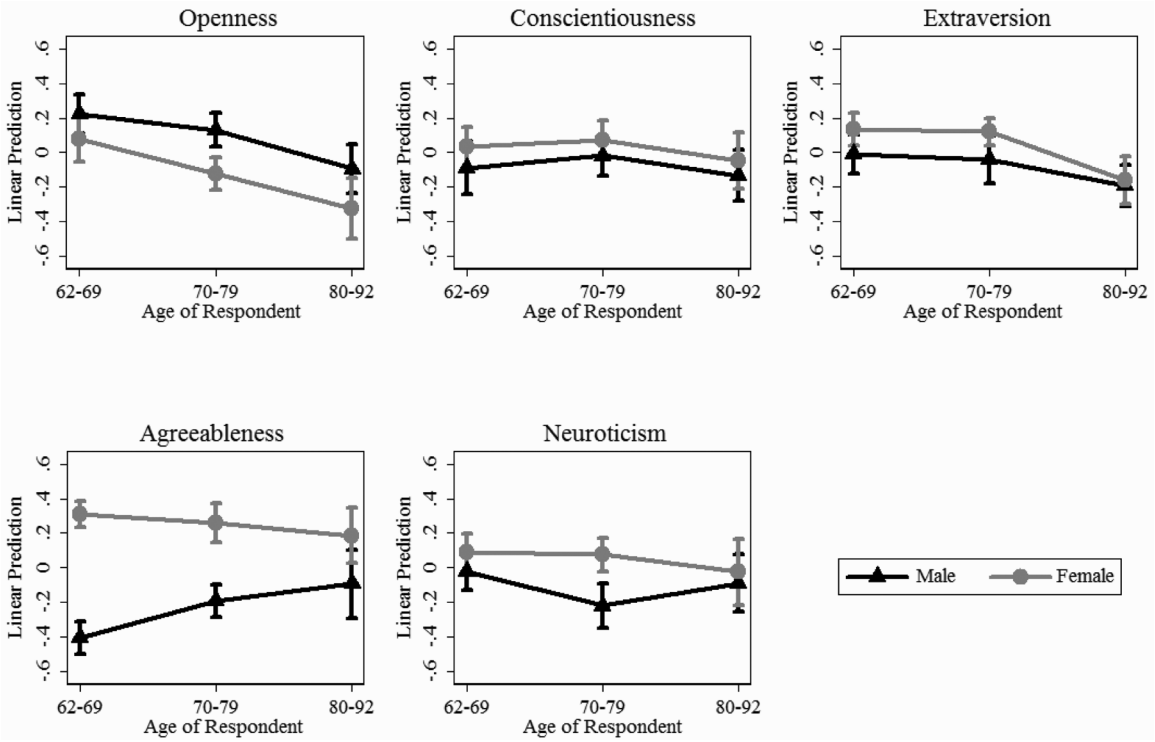

Figure 1. Age trends in personality factor scores in NSHAP W2, split by gender. NSHAP, National Social Life, Health, and Aging Project.

a more sophisticated form of SEM in order to account for the high intercorrelations among the four, positive-sounding personality traits, such as the use of a general factor (Schermer & Vernon, 2010), second-order factors (Digman, 1997), or a bifactor model (Holzinger & Swineford, 1937; Jennrich & Bentler, 2011), on the assumption that the common variance among the OCEA factors is accounted for by some additional factor or factors. Regardless, whenever possible, it seems preferable to use factor scores.

Age and Gender

A great deal of research on personality is focused on the question of age differences in personality, and as such, it provides us with heavily replicated findings concerning age trajectories in personality traits (Roberts, Walton, & Viechtbauer, 2006; Soto, John, Gosling, & Potter, 2011; Srivastava & John, 2003). In older adulthood, individuals should be largely stable in terms of their personality traits (Roberts & DelVecchio, 2000), although some evidence

suggests that the scale values that people report on all five personality traits may drop in later life (Soto et al., 2011).

Figure 1 shows line graphs which plot the mean factor scores by age category and gender. Lines were produced using the command in Stata. Lines also possess 95% confidence interval (CI) bars, in order to call attention to gender differences that are significant at $p < .05$. Point estimates were calculated using NSHAP W2 sampling weights, and CI bars were computed using settings within Stata, using the command to obtain correct variance estimates for the individuals who returned the leave-behind questionnaire. Individuals who fell outside the W2 returning respondents' age range—that is, 64–90—were excluded from the analysis. Otherwise, all respondents were included, meaning these figures do include partners. Finally, note that the y-axes all have the same range, in order to further facilitate comparisons between these graphs. Recall that the factor scores are standardized, and as such are centered on zero for the whole sample of men and women at every age. Also, note that these are estimated factor scores, since the true values of the latent construct are unknown.

Numerous personality traits do show age-related declines, although the change is most dramatic for Openness. Men and women also generally have similar levels of the Big Five traits over the age range in NSHAP, with one important exception: women were generally much more agreeable than men, and in the youngest age group, their point estimate for Agreeableness is more than a full standard deviation higher than for men. This gap begins to close in the middle age category, and for the oldest old in the NSHAP sample, there is no longer a significant difference. Both men and women contribute to closing this gap, although men more so. This pattern of convergence seems to be unique to Agreeableness in the sample. Other traits, such as Openness and Neuroticism, seem to show the widest gender gaps in the middle age category. For Conscientiousness and Extraversion, the gender gap never appears significant.

Discussion

Several patterns emerged in our analyses. We found that women generally had higher Agreeableness than men, which is commensurate with previous studies that have shown that women generally score higher on Agreeableness (Costa, Jr., Terracciano, & McCrae, 2001; Feingold, 1994). We also described age trends in Agreeableness, Openness, Extraversion, and Conscientiousness. However, we cannot use these results to say that any of these traits change with age. The MIDI battery was added to the second wave of the NSHAP survey, and we are not able to compare the results of this analysis to results from W1. The language above has tried to refrain, therefore, from discussing age differences in terms of age-related changes. Without longitudinal data on these measures, we cannot separate cohort, period, and age-related effects. The patterns that we describe in Agreeableness and Openness are, however, in accord with studies describing the relationship between personality traits and age (Roberts et al., 2006; Specht et al., 2011). Regardless, it is apparent that there are age-related differences between individuals in the NSHAP sample in personality scores, especially in men's Agreeableness. The large gap between men and women on this trait suggests that gender may be an important confounder in any multivariate analysis predicting some outcome using Agreeableness. The converse may also be true: gender differences in health behaviors, attitudes, psychological well-being, and quality of life may be partially attributable to differences in Agreeableness.

Another important methodological consideration arising from our analysis is that if a researcher estimates factor scores, and then uses them as predictors in some regression analysis, then this assumes that the scores are produced without error, and that the score produced by the factor analysis is an accurate depiction of the individual's position in the trait space. Simultaneously, estimating latent traits and predicting outcomes using those traits (i.e., fitting a SEM) would be a solution to this issue. Scales, similarly, should be approached with caution, considering the alpha values of the scales are modest, implying that they are less reliable than factor scores in proxying the underlying traits; the considerable number of cross-loadings are likely responsible for these low alpha values, since the true trait is seemingly not a function of the theoretically specified adjectives alone (e.g., both Extraversion and Agreeableness affect how highly one rates oneself on the adjective "warm"). Alternative modeling strategies such as bifactor and general factor models, may provide additional information about the properties of the MIDI battery, as well as substantial insights into personality processes and structure. In any event, it seems preferable to use factor scores instead of scales.

Conclusions

In this article, we offered recommendations for the use of the MIDI in multivariate analyses using NSHAP. We concluded that factor scores rather than averaged scales would be more appropriate for such analyses, for several reasons. First, the MIDI factor structure contains a considerable number of cross-loadings, and without taking these cross-loadings into account, model fit suffers considerably; this corresponds with findings from a similar analysis carried out on the MIDI in MIDUS (Zimprich et al., 2012). Secondly, fitting a model to the data using WLSMV allows for less missing data, meaning multivariate analyses using factor scores will not be as reduced in sample size, compared with analyses that used averaged scales. Third, regressions predicting some consequential outcomes could be combined with an ESEM or CFA in a single model, meaning factor scores would not need to be estimated, and factors could be related to this consequential outcome directly. Fifth, there may be less attenuation when using the scores as covariates, due to the shrinkage involved

in estimating the factor scores, depending on how scores are estimated. However, both scores and scales are likely to be useful in multivariate analysis, and the range of variables contained in NSHAP offer a unique opportunity to describe, and theorize, interactions between persons and environments through the medium of personality traits.

Box 1: Summary of Guidelines for Using the MIDI Personality Battery in NSHAP

- The NSHAP version of the MIDI does capture the Big Five trait structure, and as such can be used by researchers to examine the interrelationships between traits, social context, and well-being.
- Factor scores are preferable to scales, due to the large number of cross-loadings.
- Age and gender may be important confounders in analyses predicting consequential outcomes using traits. This is likely to be especially true for Agreeableness.
- Correlations between traits and outcomes may be attenuated by moderate reliabilities for the Big Five measures
- Therefore, alternative specifications of the Big Five, such as bifactor models, may provide additional traction on theoretical problems and empirical predictions.

Funding

This work was supported by the National Institutes of Health including the National Institute on Aging (R01 AG021487), the Division of Behavioral and Social Sciences Research for the National Health, Social Life, and Aging Project (NSHAP) (R37 AG030481), and the NSHAP Wave 2 Partner Project (R01 AG033903), and by NORC, which was responsible for the data collection. Coauthor, A. Tiedt was supported by a National Institute on Aging Postdoctoral Traineeship (T32 AG00243).

Correspondence

Correspondence should be addressed to James Iveniuk, PhD, NORC at the University of Chicago, 1155 E 60th Street, 2nd Floor, Chicago, IL 60637. E-mail: iveniukj@gmail.com.

References

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling*, *16*, 397–438.

Asparouhov, T., & Muthén, B. O. (2010). Weighted least squares estimation with missing data. http://www.statmodel.com/download/GstrucMissingRevision.pdf.

Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economic Letters*, *115*, 1(11–15). doi:10.1016/j.econlet.2011.11.015

Cornwell, B., Laumann, E. O., & Schumm, L. P. (2008). The Social connectedness of older adults: A national profile. *American Sociological Review*, *73*, 185–203. doi:10.1177/000312240807300201

Cornwell, E. Y., & Waite, L. J. (2009). Social disconnectedness, perceived isolation, and health among older adults. *Journal of Health and Social Behavior*, *50*, 31–48. doi:10.1177/002214650905000103

Cornwell, E. Y., & Waite, L. J. (2012). Social network resources and management of hypertension. *Journal of Health and Social Behavior*, *53*, 215–231. doi:10.1177/0022146512446832

Costa, P. T. Jr, Terracciano, A., & McCrae, R. R. (2001). Gender differences in personality traits across cultures: Robust and surprising findings. *Journal of Personality and Social Psychology*, *81*, 322–331. doi:10.1037/0022-3514.81.2.322

Costa, P. T., & McCrae, R. R. (1992a). Four ways five factors are basic. *Personality and Individual Differences*, *13*, 653–665. doi:10.1016/0191-8869(92)90236-I

Costa, P. T., & McCrae, R. R. (1992b). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. Odessa, FL: Psychological Assessment Resources.

Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*, 417–440. doi:10.1146/annurev.ps.41.020190.002221

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, *73*, 1246–1256. doi:10.1037/0022-3514.73.6.1246

Eysenck, H. (1998 [1947]). *Dimensions of personality*. New Brunswick, NJ: Transaction.

Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, *116*, 429–456. doi:10.1037/0033-2909.116.3.429

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, *9*, 466–491. doi:10.1037/1082-989X.9.4.466

Funder, D. C. (2001). Personality. *Annual Review of Psychology*, *52*, 197–221. doi:10.1146/annurev.psych.52.1.197

Goodwin, R. D., & Gotlib, I. H. (2004). Gender differences in depression: The role of personality factors. *Psychiatry Research*, *126*, 135–142. doi:10.1016/j.psychres.2003.12.024

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, *37*, 504–528. doi:10.1016/S0092-6566(03)00046-1

Hampson, S. E. (2012). Personality processes: Mechanisms by which personality traits "get outside the skin". *Annual Review of Psychology*, *63*, 315–339. doi:10.1146/annurev-psych-120710-100419

Hampson, S. E., Goldberg, L. R., Vogt, T. M., & Dubanoski, J. P. (2007). Mechanisms by which childhood personality traits influence adult health status: Educational attainment and healthy behaviors. *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association*, *26*, 121–125. doi:10.1037/0278-6133.26.1.121

Holzinger, K. J., & Swineford, F. (1937). The bi-factor method. *Psychometrika*, *2*, 41–54. doi:10.1007/BF02287965

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika*, *76*, 537–549. doi:10.1007/s11336-011-9218-4

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, third edition: Theory and research* (pp. 114–158). New York, NY: The Guilford Press.

Lachman, M. E., & Weaver, S. L. (1997). The Midlife Development Inventory (MIDI) personality scales: Scale construction and scoring. Technical Report. Waltham, MA: Brandeis University. doi:10.1177/1359105306066610

Lindau, S. T., Schumm, L. P., Laumann, E. O., Levinson, W., O'Muircheartaigh, C. A., & Waite, L. J. (2007). A study of sexuality and health among older adults in the United States. *The New England journal of medicine*, *357*, 762–774. doi:10.1056/NEJMoa067423

Lindau, S. T., Laumann, E. O., Levinson, W., & Waite, L. J. (2003). Synthesis of scientific disciplines in pursuit of health: The Interactive Biopsychosocial Model. *Perspectives in Biology and Medicine*, *46*, S74–S86.

Lodi-Smith, J., & Roberts, B. W. (2007). Social investment and personality: A meta-analysis of the relationship of personality traits to investment in work, family, religion, and volunteerism. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*, *11*, 68–86. doi:10.1177/1088868306294590

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J., Trautwein, U., & Nagengast, B. (2010). A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*, *22*, 471–491. doi:10.1037/a0019227

Marsh, H. W., Nagengast, B., & Morin, A. J. (2013). Measurement invariance of big-five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology*, *49*, 1194–1218. doi:10.1037/a0026913

Mitka, M. (2001). Aging patients are advised "stay active to stay alert". *JAMA*, *285*, 2437–2438. doi:10.1001/jama.285.19.2437-JMN0516-2-1

Muthén, L. K., & Muthén, B. O. (1998–2011). *Mplus user's guide. Sixth Edition*. Los Angeles, CA: Muthén and Muthén.

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*, 401–421. doi:10.1146/annurev.psych.57.102904.190127

Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, *17*, 354–373. doi:10.1037/a0029315

Roberts, B. W. (2007). Contextualizing personality psychology. *Journal of Personality*, *75*, 1071–1081. doi:10.1111/j.1467-6494.2007.00467.x

Roberts, B. W. (2009). Back to the future: Personality and assessment and personality development. *Journal of Research in Personality*, *43*, 137–145. doi:10.1016/j.jrp.2008.12.015

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*, 3–25. doi:10.1037/0033-2909.126.1.3

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (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

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*, 1–25. doi:10.1037/0033-2909.132.1.1

Schermer, J. A., & Vernon, P. A. (2010). The correlation between general intelligence (g), a general factor of personality (GFP) and social desirability. *Personality and Individual Differences*, *48*, 187–189. doi:10.1016/j.paid.2009.10.003

Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martínez, V. (2007). The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, *38*, 173–212. doi:10.1177/0022022106297299

Shannahan, M. J., Hill, P., Roberts, B. W., & Eccles, J. (2012). Conscientiousness and health across the life course: A critical review and new hypotheses. Paper presented at the The Working Group on Conscientiousness and Healthy Aging.

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*, 330–348. doi:10.1037/a0021717

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*, 862–882. doi:10.1037/a0024950

Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology*, *84*, 1041–1053. doi:10.1037/0022-3514.84.5.1041

Ullman, J. B., & Bentler, P. M. (2003). Structural equation modeling. In J. A. Schinka, W. F. Velicer, & I. B. Weiner (Eds.), *Handbook of psychology: Research methods in psychology*. Hoboken, NJ: John Wiley and Sons. doi:10.1002/0471264385.wei0224

Zimprich, D., Allemand, M., & Lachman, M. E. (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