Published in final edited form as:
Psychol Aging. 2008 September ; 23(3): 558–566. doi:10.1037/a0012897.

Age Differences in the Big Five Across the Life Span: Evidence from Two National Samples

M. Brent Donnellan and
Michigan State University

Richard E. Lucas
Michigan State University & German Institute for Economic Research

Abstract

Cross-sectional age differences in the Big Five personality traits were investigated using two large datasets from Britian and Germany, the British Household Panel Study (BHPS; N ≥ 14,039) and the German Socio-Economic Panel Study (GSEOP; N ≥ 20,852). Participants ranged in age from 16 to the mid 80s and completed a 15-item version of the Big Five Inventory (e.g., John & Srivastava, 1999) in either 2005 or 2006. The observed age trends were generally consistent across both datasets. Extraversion and Openness were negatively associated with age whereas Agreeableness was positively associated with age. Average levels of Conscientiousness were highest for participants in middle age. The one exception was that Neuroticism was slightly negatively associated with age in the BHPS and slightly positively associated with age in the GSEOP. Neither gender nor education level were consistent moderators of age differences in the Big Five.

Keywords

Personality Assessment; Big Five; Personality Development; British Household Panel Study; German Socio-Economic Panel Study

Age-related differences in personality have captured human attention for centuries. For instance, Aristotle devoted three chapters of Book II of his Rhetoric to the description of the characteristics of individuals at different phases of the life span. His prediction, translated into a testable hypothesis, is that there should be age-related differences in personality attributes. Conversely, William James (1892/1985) believed that character was fixed by age 30 so that there would be little reason to expect age-linked personality differences after this point in the life span (see Kelly, 1955). In contemporary psychology, much of the interest in this question has fomented around Costa and McCrae's work on personality in adulthood (Helson, Kwan, John, & Jones, 2002; McCrae & Costa, 2003; Roberts, Walton, & Viechtbauer, 2006; Roberts, Wood, & Smith, 2005; Srivastava, John, Gosling, & Potter, 2003; Terracciano, McCrae, Brant, & Costa, 2005). The present analyses contribute to this literature by evaluating cross-sectional age differences in the Big Five personality traits in two large national datasets.

Address correspondence to: M. Brent Donnellan Department of Psychology Michigan State University East Lansing, MI 48823 E-mail: donnel59@msu.edu. Correspondence regarding this manuscript should be addressed to M. Brent Donnellan, Department of Psychology, Michigan State University, East Lansing, MI 488233. Email: donnel59@msu.edu..

Age Differences in the Big Five

There is something of a consensus that five broad domains capture much of the variability in personality traits (John & Srivastava, 1999; but see Ashton & Lee, 2007; Block, 1995). These "Big Five" are Extraversion (traits like energetic and sociable), Agreeableness (traits like considerate and kind), Conscientiousness (traits like hard-working and orderly), Neuroticism (traits like nervous and tense), and Openness (traits like artistic and creative). Mean-level differences in the Big Five across the life span were summarized by Roberts et al. (2006) who compiled the results of 113 longitudinal samples involving 50,120 participants. Different longitudinal studies examined different personality traits for varying periods of the life span so the degree of comprehensiveness varied for each of the traits they examined. Moreover, Roberts and colleagues drew on a distinction between two aspects of Extraversion, traits related to independence and dominance (labeled Social Dominance) versus traits related to positive affect, activity level, and sociability (labeled Social Vitality) following Helson and Kwan (2000).

Roberts et al. (2006) found that average levels of Social Vitality tended to be fairly stable across the life span, although there was a slight spike from adolescence to the early 20s followed by mean-level consistency from the mid 20s until the mid 50s when there was a slight decline. In contrast, Social Dominance showed a more pronounced and consistent increase from adolescence to the mid 30s when mean-levels remained consistent until the mid 50s. Data on average levels of this trait were not available beyond this point in the life span because only 7 studies examined changes in Social Dominance for participants in their 50s and older. Agreeableness and Conscientiousness showed relatively gradual increases in absolute scores across the life span whereas Neuroticism showed relatively gradual decreases. Lastly, Openness showed a mean-level increase from adolescence to the early 20s and then mean-levels remained fairly consistent until the mid 50s when average levels started to decline.

Terracciano, McCrae, Brant, and Costa (2005) conducted cross-sectional and longitudinal analyses examining links between age and mean-levels of the Big Five. This study was completed too recently to be included in the Roberts et al. (2006) meta-analysis. Terracciano et al. found that scores on Extraversion generally declined from age 30 to 90 although the drop in Extraversion was more pronounced after the mid 50s or so. Agreeableness demonstrated a fairly linear increase with age whereas the pattern for Conscientiousness was curvilinear: scores increased up to a peak somewhere between the ages of 50 to 70 and then declined. Average levels of Neuroticism generally declined with age but increased slightly starting around age 80. Finally, Openness showed a negative and linear association with age. In general, Terracciano found similar cross-sectional and longitudinal results with the exception that the cross-sectional zenith for Conscientiousness was around age 50 as compared to around age 70 for the longitudinal analyses.

Age differences in the Big Five have also been identified in cross-cultural research. McCrae et al. (1999) used convenience samples from Germany, Italy, Portugal, Croatia, and Korea and found that self-reports of Extraversion and Openness were lower in older participants than younger participants whereas Conscientiousness and Agreeableness showed the reverse pattern. Results were mixed for Neuroticism as it was found to be lower in older participants versus younger participants in Germany, Portugal, and Korea whereas age differences were not statistically detectable in Italy and Croatia. McCrae et al. (2005) found that observers rated adults (ages 4−98) higher on measures of Conscientiousness but lower on measures of Extraversion and Openness when contrasted with college students in research that included participants from 50 countries. Age differences for Neuroticism and Agreeableness were moderated by gender: the negative association between age and observer reports of

Neuroticism was more evident for men than women whereas the positive association between age and observer reports Agreeableness was more evident for women than men.

In sum, the existing evidence broadly suggests that levels of Agreeableness and Conscientiousness are positively associated with age whereas levels of Extraversion and Openness are negatively associated with age (see also Costa, McCrae, Zonderman, Barbano, Lebowitz, & Larson 1986; Helson et al., 2002; Mroczek, Spiro, & Griffin, 2006; Srivastava et al., 2003). Average levels of Neuroticism are generally negatively associated with age although there are exceptions in the literature such that consistent age differences have not emerged in all countries and there are hints that this trait may increase around age 80. All in all, these broad trends translate to age-linked increases on Digman's (1997) alpha factor and age-linked decreases on Digman's beta factor. In other words, as individuals grow older, they seem to increase on traits related to social interest and communion and decrease on traits related to agency and zestful approach to life.

Despite the emergence of some relatively consistent age differences in the Big Five, there are a couple of limitations of the existing literature that are worth noting. First, no single study has examined age differences in all of the Big Five using nationally representative samples. To be sure, Costa et al. (1986) examined cross-sectional age differences in Neuroticism, Extraversion, and Openness in a representative sample from the United States and Steunenberg, Twisk, Beekman, Deeg, and Kerkhof (2005) examined changes in Neuroticism in a representative sample from The Netherlands; however, neither study examined all five traits simultaneously. Indeed, the absence of data from nationally representative samples is perhaps the most serious limitation of this entire literature (see Roberts et al., 2006, p. 20). For example, most of the participants in Terracciano et al. (2005) were described as “generally healthy and highly educated” (p. 494). Second, few studies include participants past the age of 70 (Terracciano et al., 2005).

The goal of this report is extend research on age differences in the Big Five using data from two large national panel studies that include participants past the age of 70. In addition, we explore whether or not gender and education moderates cross-sectional age differences. Although there is persistent interest in gender differences in mean-levels of personality (e.g., Chapman, Duberstein, Sörensen, & Lyness, 2007; Costa, Terracciano, & McCrae, 2001; Feingold, 1994; Goldberg, Sweeney, Merenda, & Hughes, 1998), evidence that gender moderates age differences in the Big Five is not generally robust (see Roberts et al. 2006). Likewise, the existing literature does not suggest that education effects will be large (e.g. Costa et al., 1986; Goldberg et al., 1998; Löckenhoff et al., in press). Thus, we view these analyses for gender and education as exploratory given the lack of clear trends in the existing literature.

Method

Samples

The British Household Panel Study (BHPS) is an on-going panel study of British households that began in 1991. Households were selected using a multi-stage probability design with systematic sampling, and all members of the household ages 16 and older were asked to participate. The sample was initially representative of the population of the United Kingdom, though as would be expected, some attrition has occurred. Average annual attrition is relatively low, with about 5% of initial participants lost to attrition each year. Additional sub-samples were recruited after 1991, and the cumulative attrition at the time of the Big Five assessment for these samples is thus lower than for the original sample.

The age range of the most recent wave of BHPS was 16 to 99 years but we trimmed the sample to participants who were less than 86 years old because of the sparse number of participants

(defined as $n < 40$) at older ages. The BHPS sample was approximately 54% women and the sample sizes used in these analyses ranged from 14,039 to 14,055 depending on the Big Five trait. The average age of the sample with complete Big Five data was 45.29 years (SD = 18.04). Big Five measures were administered to the panel at the most recent wave (Wave 15) which was predominately collected in September through December of 2005 (approximately 94% of all Wave 15 participants). The rest of the data were collected before May of 2006. Participants completed portions of the survey in face-to-face interviews, though the Big Five measures were administered in a self-completion format.

The German Socio-Economic Panel Study (GSOEP) is an on-going study of German households that began in 1984 (see Haisken-De New & Frick, 2005, for details). Households were selected using a multi-stage random sampling technique, and all members of the household ages 16 and older were asked to participate. Like the BHPS, multiple samples were recruited in the years since the study began, and each sample was initially representative of the population from which it was drawn. Again, however, some attrition has occurred, which raises questions about the representativeness of the remaining sample. Attrition rates for the earliest sample are similar to those from the BHPS, with average yearly attrition at around 6%. Again, it is important to note that although the cumulative attrition for the original sample is substantial (which might affect the representativeness), there have been a number of sub-samples added to the study much more recently. For instance, 44% of the sample that was used in this analysis was recruited in 2000 or 2002. These sub-samples have much lower rates of cumulative attrition because they have been in the study for much less time than the original sample.

Big Five measures were administered to the panel at the most recent wave which was collected in 2005. The age range of the most recent wave of the GSOEP was 16 to 95 but we trimmed the sample to those participants who were less than 85 given the sparse number of participants past this cut-off (i.e., $n < 40$). The GSOEP sample was approximately 52% women and the sample sizes used in these analyses ranged from 20,852 to 20,876 depending on the Big Five trait. The average age of the sample with complete Big Five data was 46.03 years (SD = 17.23). Participants completed the measure through an oral interview (roughly 26%), by a written questionnaire (roughly 50%), or by Computer Assisted Personal Interviewing (roughly 25%).[1]

Personality

Participants completed a 15-item version of the Big Five Inventory (BFI; John & Srivastava, 1999) using a seven point scale (BHPS: 1 = “Does not Apply” to 7 = “Applies perfectly,” GSOEP: 1 = “Does not apply” to 7 = “Does Apply”). Three items were used to measure each dimension.[2] Internal consistencies in the BHPS were as follows: Extraversion (alpha = .54, average inter-item r = .28), Agreeableness (alpha = .53, average inter-item r = .28), Conscientiousness (alpha = .52, average inter-item r = .28), Neuroticism (alpha = .68, average

[1]We examined whether format type was associated with scores on the Big Five. It appeared that scores derived from both oral interviews and CAPI formats were higher than written formats for some traits. Accordingly, we collapsed interviews and CAPI administrations to contrast with written administrations. The main effects of format type in the d-metric were as follows: Extraversion: d = .03; Agreeableness: d = .20; Conscientiousness: d = .24; Neuroticism: d = −.14; Openness: d = −.01 (positive values indicated interview formats were higher). It was also the case that older participants were more likely to use interview based formats versus written formats when compared to younger participants. The real question, however, was whether or not format type moderated age trends. We first used an ANCOVA model to control for format type (1 = interview or CAPI, 0 = written) for the means reported in Table 1. The predicted means were quite similar to those reported in Table 1 which suggests that differences in format type did not create major confounds (Table available upon request). We also conducted tests of interactions using same regression strategy that we used for gender and education (see Results). Only one instance met our threshold for a meaningful interaction – Conscientiousness. It appeared that age differences in Conscientiousness after mid-life were more pronounced for written administrations than for interview-based administrations. Nonetheless, the shape of the predicted age curves was quite similar.

[2]Extraversion was assessed with the BFI items 1, 6, and 36. Agreeableness was assessed with items 17, 32, and 37. Conscientiousness was assessed with items 3, 23, and 33. Neuroticism was assessed with items 9, 19, and 39. Openness was assessed with items 5, 20, and 30.

inter-item *r* = .41), and Openness (alpha = .67, average inter-item *r* = .41). Internal consistencies in the GSOEP were as follows: Extraversion (alpha = .66, average inter-item *r* = .41), Agreeableness (alpha = .51, average inter-item *r* = .28), Conscientiousness (alpha = .62, average inter-item *r* = .39), Neuroticism (alpha = .60, average inter-item *r* = .33), and Openness (alpha = .63, average inter-item *r* = .37).

To establish the utility of these short scales, we correlated these short scales with the full BFI scales using data from the Gosling-Potter Internet Personality Project (N = 628,640; see Srivastava et al., 2003). Our 3-item scales were strongly correlated with the full versions for all five BFI scales (Extraversion: .90; Agreeableness: .88; Conscientiousness: .88; Neuroticism: .89; Openness: .86) and our 3-items scales were strongly correlated with the remaining five to seven items in each BFI scale that were not included in our short measure (Extraversion: .73; Agreeableness: .71; Conscientiousness: .73; Neuroticism: .70; Openness: .70) Based on these results, we had confidence that these brief measures were reasonable substitutes for the longer scales. For additional evidence about the utility of shorter forms of longer Big Five assessments see Donnellan, Oswald, Baird, and Lucas (2006) and Rammstedt and John (2007).

Mean differences between the samples might stem from both artifacts (e.g., differences in the response scales and item translations) as well as real cross-cultural differences in personality. It is difficult to disentangle these issues with the present data and these concerns are not the focus of the present research. Accordingly, we calculated *T* scores within each sample to control for any mean-level differences across the two samples following the logic used by McCrae et al. (1999). Scores were norm referenced to the group of participants aged 30 to 34 within each sample which facilitates an intuitive within sample comparison: scores below 50 indicate trait levels that are lower than the reference group whereas scores above 50 indicate trait levels that are higher than the reference group. Such differences can be interpreted against the overall *T* score SD of 10 for rough effect size calculations. For example, the average score on Conscientiousness for individuals ages 16 to 19 was 42.76 (SD = 10.97) in the BHPS and 41.49 (SD = 12.27) in the GSOEP (see Table 1). Both of these values indicate that late adolescents scored more than 7 tenths of a standard deviation lower than individuals in their early 30s, a pattern which is consistent with longitudinal research that has found increases on traits linked with Conscientiousness during the transition from adolescence to adulthood (e.g., Caspi, Roberts, & Shiner, 2005; Donnellan, Conger, & Burzette, 2007).

Education

We coded whether participants in each study had completed basic education in their countries of origin. Following Hu, Stewart-Brown, Twigg, and Weich (2007) we used whether or not participants had educational qualifications in the BHPS for this purpose and we used whether or not participants completed the equivalent of high school in the GSOEP using data from the Cross-National Equivalent File, a dataset constructed from the GSOEP responses to facilitate cross-national comparisons across five panel studies conducted in the U.S., U.K., Canada, Germany, and Australia. We restricted the examination of education related differences to those participants who were 30 years or older to avoid confusing education-linked differences with age-linked differences in personality. Approximately 23% of those 30 years and older did not have any qualifications in the BHPS whereas approximately 14% of those 30 years or older did not complete the equivalent of high school in the GSOEP.

Results

Overview of Analytic Strategy

We first created 8 age groups (16 to 19 year olds, 20 to 29 year olds, 30 to 39 year olds, 40 to 49 year olds, 50 to 59 year olds, 60 to 69 year olds, 70 to 79 year olds, and individuals over 80) to examine broad age trends in the Big Five. We then used a hierarchical regression approach to formally model the association between age and personality traits. Given the large sample sizes, we placed a higher premium on the size of the effects and on obtaining replicable models rather than statistical significance. This follows in the tradition of the "less is more" perspective used by McCrae et al. (1999) and Srivastava et al. (2003). The most complicated models that we evaluated involved cubic terms for age, a decision motivated by an inspection of the age curves (see Figure 1) and the fact that few other studies have used anything higher than cubic terms to model age-personality curves (e.g., Srivastava et al., 2003; Terracciano et al., 2005). Following Srivastava et al. (2003), we selected more complicated models (e.g., models with quadratic terms for age as opposed to simpler linear models) when the higher order term improved overall model fit at $F > 25.00$. We followed similar strategies for addressing questions about the potential moderating effects of gender and education.

Age Differences in the Big Five from Ages 16 to the Mid 80s

Table 1 displays average levels of personality traits for the 8 age groups. To facilitate an intuitive understanding of the direction and magnitude of the age differences, we identified the age group associated with the maximum score for each trait and the age group associated with the minimum score for each trait. We then computed the simple difference between the two groups which yields a difference score that is expressed in age 30−34 *T* "units." We considered differences of around 2 points as "small" differences, differences of around 5 points as "medium" differences, and differences above 8 as "large" – designations that parallel the conventions used for interpreting standardized mean differences (e.g., McCartney & Rosenthal, 2000).

We first examined the pattern of age differences in the BHPS. Extraversion showed a linear pattern of age differences such that the youngest group scored the highest and the oldest group scored the lowest on this measure. This difference was nearly "large" (7.60 *T* units). Agreeableness showed the opposite linear pattern such that the oldest group scored highest and the youngest group scored lowest, however the difference was fairly "small" (2.83 *T* units). Conscientiousness showed a curvilinear pattern of age differences such that there was a "large" difference (8.06 *T* units) between the youngest age group and the 40−49 year olds (the highest group) whereas there was a "small" to "medium" difference between the 40−49 year olds and the oldest age group (4.06 *T* units). Neuroticism demonstrated a linear pattern of age differences whereby the youngest group scored the highest and the 70−79 year olds scored the lowest (4.22 *T* units). Finally, Openness showed a fairly linear pattern such that the 20−29 year olds scored the highest and the oldest group scored the lowest; the difference was "large" by our conventions (8.61 *T* units).

Roughly similar patterns of age differences were observed in the GSOEP. The youngest group scored the highest and the 70−79 year old group scored the lowest for Extraversion and the difference was "small" to "medium" (3.64 *T* units). The oldest group scored highest and the youngest group scored lowest for Agreeableness (a difference of 4.52 *T* units). There was a "large" difference (9.73 *T* units) between the youngest age group and the 40−49 year olds for Conscientiousness whereas there was a slight difference between the 40−49 year old group and the oldest group (1.38 *T* units). Neuroticism showed a fairly linear increase in Germany such that the youngest group scored the lowest and the 60−69 year old group scored the highest. This difference was "small" (2.71 *T* units). Last, Openness showed a linear pattern of age

differences whereby the 16−19 year olds scored the highest and the oldest group scored the lowest group; the difference was "medium" by our conventions (6.19 *T* units). All and all, we concluded that the most notable difference between the BHPS and GSOEP were that that Neuroticism trends were in the opposite directions. Even so, there was a considerable degree of overlap in the broad patterns of age differences across the two datasets. To be sure, Conscientiousness appeared to be the personality domain with the largest replicable effect size when considering the difference between adolescent and middle age participants.

We then modeled age differences using regression analyses and results are reported in Table 2 and displayed in Figure 1. Age was mean-centered within each dataset before higher order terms were calculated. The regression results mapped closely to the age trends depicted in Table 1. That is, Extraversion and Openness scores were predicted to be lower in older individuals as compared to younger individuals whereas scores on Agreeableness were predicted to be higher in older individuals compared to younger individuals. Age differences in Conscientiousness had a predicted curvilinear pattern such that average levels of this trait were highest in middle adulthood. Last, age difference in N**e**uroticism diverged in the two samples along the previously described lines - older individuals were predicted to score lower on Neuroticism than younger individuals in the BHPS whereas this pattern was reversed in the GSOEP.

Gender and Education Effects

We first examined overall gender differences in the Big Five using *d* metric effect sizes which were scored so that positive scores indicated that women scored higher than men. These are displayed in Table 3 and the effect sizes were generally similar in the BHPS and the GSOEP. The one caveat was that the direction of the gender difference for Openness was reversed across the two datasets. In addition, the overall gender differences were more or less consistent with the effect sizes reported by Costa et al. (2001) given that we found the biggest differences for Neuroticism, Agreeableness, and Extraversion and relatively smaller effect sizes for Conscientiousness and Openness. We also calculated *d* metric effect sizes within each of our 8 age groups and these differences are also reported in Table 3. As seen in Table 3, there did not seem to be striking evidence that effect sizes varied systematically across the age groups in either dataset.

We then formally evaluated whether gender moderated the association between age and personality. For these analyses, we followed a hierarchical strategy similar to the strategy we used to model higher order terms involving non-linear age effects. We tested for interactions involving gender by modeling effects in six steps: 1) a model with gender and linear age terms; 2) a model with gender, linear age, and gender by linear age terms; 3) a model with gender, linear age, gender by linear age, and quadratic age terms; 4) a model with gender, linear age, gender by linear age, quadratic age, and gender by quadratic age terms; 5) a model with gender, linear age, gender by linear age, quadratic age, gender by quadratic age, and cubic age terms; and finally 6) a model that added the gender by cubic age term to the previous model. We were interested in whether the gender by age interaction terms improved overall model fit at $F > 25.00$ for Models 2, 4, and 6 compared with Models 1, 3, and 5, respectively. None of the F change statistics for the relevant comparisons suggested that gender acted as a moderator of age effects according to our standards. This was the case for both the BHPS and the GSEOP datasets. These results seem consistent with the conclusion drawn by Roberts et al. (2006) that "there is very little support for the idea that men and women change in distinct ways" (p. 15).

We followed a similar approach to examine questions about education. Recall that these analyses were restricted to individuals who were 30 years or older to cover the era in the life span when most people would have completed formal schooling in both countries. Effect sizes were calculated so that positive scores indicated that individuals with more education scored

higher than individuals with less education. These effect sizes were roughly similar in the two datasets as seen in Table 3 and the largest overall effect sizes were for Openness. As a point of comparison, we calculated *d* metric effect sizes from Table 1 in Löckenhoff et al. (in press) who compared individuals who completed more than 12 years of education to those who did not complete at least 12 years of education in a sample from the United States. The *d* metric effect sizes in that report were .09, −.13, .20, −.28, and .57 for Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness, respectively. Thus, Löckenhoff et al. (in press) found the most pronounced difference for Openness which was consistent with our results.

We also calculated *d* metric effect sizes within each of our age groups (see Table 3) and the education effects looked generally similar across age groups in both datasets. Last, we formally evaluated whether education status moderated the association between age and personality. For these analyses, we followed the same general strategy outlined for testing for gender effects (i.e. we replaced gender and gender by age interaction terms with education and education by age interaction terms). None of the relevant model comparisons suggested that education acted as a compelling moderator of age differences. This was the case in both the BHPS and the GSEOP datasets.

General Discussion

To our knowledge, this is the first study that examined age differences in all of the Big Five from ages 16 to the mid 80s using two large, national datasets from Britain and Germany. In both countries, there is evidence that Extraversion and Openness are negatively associated with age whereas Agreeableness is positively associated with age. Likewise, mean levels of Conscientiousness were highest for middle age participants in both Britain and Germany. In general, cross-sectional age differences in the Big Five were detectable past age 30 which seems inconsistent with the “hard” plaster hypothesis or the idea that personality traits are completely fixed at some point in the life span (see also Srivastava et al., 2003). In addition, there did not appear to be consistent evidence that gender or education level moderated cross-sectional age differences. We now comment on several of the more notable results.

One of our more interesting findings concerned the curvilinear association between age and Conscientiousness. The association between age and Conscientiousness is often broadly characterized as linear in the existing literature. However, Terracciano et al. (2005) also reported a curvilinear association between age and this trait. We also found that scores on Conscientiousness had a curvilinear association with age such that the biggest age differences were found when comparing average levels for late adolescents with average levels of middle-aged participants. It even appeared that average levels were lower in the oldest adults when compared to middle-aged participants. One explanation for any apparent discrepancies between these results for Conscientiousness and the existing literature is that few studies have actually examined this association past age 60 using analytic strategies that are capable of detecting nonlinear effects. That is, Srivastava et al. (2003) did not include adults older than 60 in their study and the strategy of comparing groups of older and younger adults (e.g., McCrae et al., 1999) or older and very old adults (e.g., Weiss et al., 2005) does not permit an evaluation of non-linear associations. Thus, our results fill an important gap in the existing literature and point to a more nuanced association between age and Conscientiousness. Future work is needed to resolve the discrepancies between the cross-sectional results and the longitudinal results of Terracciano et al. (2005) showing that the peak average level of Conscientiousness was near age 70.

Given that we found replicable age-linked differences for four out of the five Big Five traits, a natural question is why such age differences exist. There are currently two dominant

explanations for age differences in personality traits -- the intrinsic maturation perspective and the life course perspective. The intrinsic maturational argument holds that normative age-related changes in personality adulthood are driven by "preprogrammed" biological processes (e.g., Costa & McCrae, 2006) whereas the life course argument posits that the major roles of adult life involving occupational pursuits, romantic relations, and parenthood drive adult personality development (e.g., Helson et al., 2002; Roberts et al., 2005). As we have argued elsewhere (Donnellan et al., 2007), it is ultimately difficult to conduct crucial tests of these two explanations because true experiments are neither feasible nor ethical in the realm of adult personality development.

Fortunately, there are alternatives to true experiments for partially resolving this debate. Evidence of replicable associations between role changes and personality changes would favor the life course perspective over a strict intrinsic maturation explanation. However, there is much controversy over the conclusiveness of the existing evidence for associations between personality changes and adult life experiences or role transitions (see e.g., the exchange between Costa & McCrae, 2006 and Roberts et al., 2006b). Thus, more work is needed and the general interest in resolving this debate may be helpful for convincing those in charge of the BHPS and GSOEP to administer measures of the Big Five in future assessments. As it stands, the important contribution of the present study are that we have presented clear evidence that (a) age differences in the Big Five are detectable in large national datasets; (b) the pattern of age differences are similar in Britain and Germany with the noteworthy exception of Neuroticism; and (c) neither gender nor education seem to consistently moderate cross-sectional age differences across the two datasets.

Puzzles, Limitations, and Conclusions

This investigation yielded at least one puzzling finding – namely the inconsistent results for Neuroticism in Britain and in Germany. In particular, the results for the German sample were somewhat exceptional given that older individuals were found to score higher on this trait than younger individuals. On the one hand, the absolute difference between the youngest group and the oldest group was very small when considered in *T* units. On the other hand, the general trend fails to replicate the broad trends in the existing literature. Future cross-sectional work in Germany may benefit from using a longer measure of Neuroticism that assesses its lower order facets to provide a more nuanced understanding of this issue. Moreover, future work using samples from other nations is needed to examine other potential cross-national differences in the association between age and Neuroticism.

In addition to this puzzle, there are at least three important limitations of the present work. The first has to do with the "representativeness" of the oldest members of the samples. The issue is that the oldest individuals in the BHPS and GSOEP are actually select members of their respective birth cohorts given that they have lived longer than what would be considered typical (see Hofer & Sliwinski, 2006). Such a select sub-sample may not be ideal for drawing inference about normative levels of traits for older individuals. This fact is one of the major conceptual and methodological issues facing researchers interested in aging and personality – the oldest participants (e.g., those in their 70s) in current samples are somewhat exceptional in terms of their longevity and it is possible that they may be exceptional in terms of their personalities.

The second limitation is the relatively low internal consistencies of our measures. In our favor, the items on these short scales were drawn from a reliable and well-validated parent instrument. Likewise, it is useful to bear in mind that there are constraints on the length of assessments in these large panel studies and extensive personality assessments are not always possible. However, it is important to be precise about the consequences of measurement error -- measurement error attenuates the ability to find differences which would limit our ability to

find systematic age trends. The fact that many of our results replicated previous research perhaps mitigates concerns over this issue.

The more serious final limitation is the cross-sectional design of this study. We readily acknowledge that cross-sectional studies are unable to disentangle age effects from cohort effects (e.g., Costa & McCrae, 1982). Future longitudinal and cohort-sequential studies are clearly needed to address these issues. On the other hand, the current evidence in favor of strong cohort effects on the Big Five is generally mixed and inconclusive (see McCrae & Costa, 2003, p. 80). For example, meta-analytic findings by Twenge (2000) suggest that there are cohort-linked increases in Neuroticism but these findings are inconclusive because they are based on convenience samples (i.e., nonprobability samples; see Donnellan & Trzesniewski, 2008) and these effects have not always replicated (e.g. McCrae & Costa, 2003; Roberts et al., 2006; Terracciano et al., 2005). Nonetheless, it is possible that different socio-historical factors in Britain and Germany might explain the divergent patterns for Neuroticism.

In closing, we believe that the present findings are noteworthy because the age trends were derived from two large panel studies and were generally consistent across the two countries. At present, we think it is safe to conclude that there are real age differences in personality and these many of these differences generalize to broad populations of individuals in Western countries. At this point, Aristotle appears to have been right and William James appears to have been wrong -- normative personality differences exist after age 30 and there is good reason to think that detectable absolute changes in personality occur across the life span.

Acknowledgements

The data used in this paper were made available by the German Institute for Economic Research and by the UK Data Archive. The BHPS data were originally collected by the ESRC Research Centre on Micro-social Change at the University of Essex, now incorporated within the Institute for Social and Economic Research. Neither the original collectors of the data nor the Archive bear any responsibility for the analyses or interpretations presented here. Preparation of this manuscript was supported by National Institute on Aging grants 1R03AG026028-01 and 1R03AG028744-01. We thank Samuel Gosling and Sanjay Srivastava for providing us with the correlations between our scales and the full BFI scales.

References

Ashton MC, Lee K. Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. Personality and Social Psychology Review 2007;11:150–166. [PubMed: 18453460]

Block J. A contrarian view of the Five-Factor approach to personality description. Psychological Bulletin 1995;117:187–215. [PubMed: 7724687]

Caspi A, Roberts BW, Shiner RL. Personality development: Stability and change. Annual Review of Psychology 2005;56:453–484.

Chapman BP, Duberstein PR, Sörensen S, Lyness JM. Gender differenes in Five Factor Model personality traits in an elderly cohort. Personality and Individual Differecnces 2007;43:1594–1603.

Costa PT Jr. McCrae RR. An approach to the attribution of aging, period, and cohort effects. Psychological Bulletin 1982;92:238–250.

Costa PT Jr. McCrae RR, Zondeerman AB, Barbano HE, Lebowitz B, Larson DM. Cross-sectional studies of personality in a national sample: 2. Stability in neuroticism, extraversion, and openness. Psyhology and Aging 1986;1:144–149.

Costa PT Jr. McCrae RR. Age changes in personality and their origins: Comment on Roberts, Walton, and Viechtbauer (2006). Psychological Bulletin 2006;132:26–28. [PubMed: 16435955]

Costa PT Jr. Terracciano A, McCrae RR. Gender differences in personality traits across cultures: Robust and surprising findings. Journal of Personality and Social Psychology 2001;81:322–331. [PubMed: 11519935]

Digman JM. Higher-order factors of the Big Five. Journal of Personality and Social Psychology 1997;73:1246–1256. [PubMed: 9418278]

Donnellan MB, Conger RD, Burzette RG. Personality development from late adolescence to young adulthood: Differential stability, normative maturity, and evidence for the maturity-stability hypothesis. Journal of Personality 2007;75:453–472.

Donnellan MB, Oswald FL, Baird BM, Lucas RE. The Mini-IPIP scales: Tiny-yet-effective measures of the big five factors of personality. Psychological Assessment 2006;18:192–203. [PubMed: 16768595]

Donnellan, MB.; Trzesniewski, KH. How should we study generational change – or should we? A critical examination of the evidence for “Generation Me.”. 2008. Unpublished manuscript

Feingold A. Gender differences in personality: A meta-analysis. Psychological Bulletin 1994;116:429–456. [PubMed: 7809307]

Goldberg LR, Sweeney D, Merenda PF, Hughes JE Jr. Demographic variables and personality: The effects of gender, age, education, and ethnic/racial status on self-descriptions of personality attributes. Personality and Individual Differences 1998;24:393–403.

Helson R, Jones C, Kwan VSY. Personality change over 40 years of adulthood: Hierarchical linear modeling analyses of two longitudinal samples. Journal of Personality and Social Psychology 2002;83:752–766. [PubMed: 12219867]

Helson, R.; Kwan, VSY. Personality development in adulthood: The broad picture and processes in one longitudinal sample.. In: Hampson, S., editor. Advances in personality psychology. 1. Routledge; London: 2000. p. 77-106.

Helson R, Kwan VSY, John OP, Jones C. The growing evidence for personality change in adulthood. Findings from research with personality inventories. Journal of Research in Personality 2002;36:287–306.

Hofer, SM.; Sliwinski, MJ. Design and analysis of longitudinal studies on aging.. In: Birren, JE.; Schaie, KW., editors. Handbook of the Psychology of Aging. 6th ed.. Academic Press; New York: 2006. p. 15-37.

Hu Y, Stewart-Brown S, Twigg L, Weich S. Can the 12-item General Health Questionnaire be used to measure positive mental health? Psychosomatic Medicine 2007;37:1005–1013.

James, W. Psychology: The briefer course. University of Notre Dame; Notre Dame, IN: 18921985.

John, OP.; Srivastava, S. The Big Five trait taxonomy: History, measurement, and theoretical perspectives.. In: Pervin, LA.; John, OP., editors. Handbook of personality: Theory and Research. 2nd ed.. Guilford Press; New York: 1999. p. 102-138.

Kelly EL. Consistency of the adult personality. American Psychologist 1955;10:659–681.

Löckenhoff CE, Terracciano A, Bienvenu OJ, Patriciu NS, Nestadt G, McCrae RR, Eaton WW, Costa PT Jr. Ethnicity, education, and the temporal stability of personality traits in the East Baltimore Epidemiologic Catchment Area study. Journal of Research in Personality. (in press)

McCartney K, Rosenthal R. Effect size, practical importance, and social policy for children. Child Development 2000;71:173–180. [PubMed: 10836571]

McCrae, RR.; Costa, PT. Personality in adulthood: A five-factor theory perspective. 2nd Edition. Guilford Press; New York: 2003.

McCrae RR, Costa PT, Pedroso de Lima, Simões A, Ostendorf F, Angleitner A, et al. Age differences in personality across the adult life span: Parallels in five cultures. Developmental Psychology 1999;35:466–477. [PubMed: 10082017]

McCrae RR, Terracciano A. Universal features of personality traits from the observer's perspective: Data from 50 cultures. Journal of Personality and Social Psychology 2005;88:547–561. [PubMed: 15740445]

Mroczek, DK.; Spiro, A., III; Griffin, PW. Personality and aging.. In: Birren, JE.; Schaie, KW., editors. Handbook of the Psychology of Aging. 6th ed.. Academic Press; New York: 2006. p. 363-377.

Rammstedt B, John OP. Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality 2007;41:203–212.

Roberts BW, Wood D, Smith JL. Evaluating the five factor theory and social investment perspective on personality trait development. Journal of Research in Personality 2005;39:166–184.

Roberts BW, Walton KE, Viechtbauer W. Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. Psychological Bulletin 2006;132:1–25. [PubMed: 16435954]

Srivastava S, John OP, Gosling SD, Potter J. Development of personality in early and middle adulthood: Set like plaster or persistent change? Journal of Personality and Social Psychology 2003;84:1041–1053. [PubMed: 12757147]

Steunenberg B, Twisk JWR, Beekman ATF, Deeg DJH, Kerkhof AJFM. Stability and change of neuroticism in aging. Journal of Gerontology: Psychological Sciences 2005;60B:P27–P33.

Terracciano A, McCrae RR, Brant LJ, Costa PT Jr. Hierarchical linear modeling analyses of the NEO-PI-R scales in the Baltimore Longitudinal Study of Aging. Psychology and Aging 2005;20:493–506. [PubMed: 16248708]

Twenge JM. The age of anxiety? Birth cohort change in anxiety and neuroticism, 1952-1993. Journal of Personality and Social Psychology 2000;79:1007–1021. [PubMed: 11138751]

Weiss A, Costa PT Jr. Karuza J, Duberstein PR, Friedman B, McCrae RR. Cross-sectional age differences in personality among Medicare patients aged 65 to 100. Psychology and Aging 2005;20:182–185. [PubMed: 15769223]

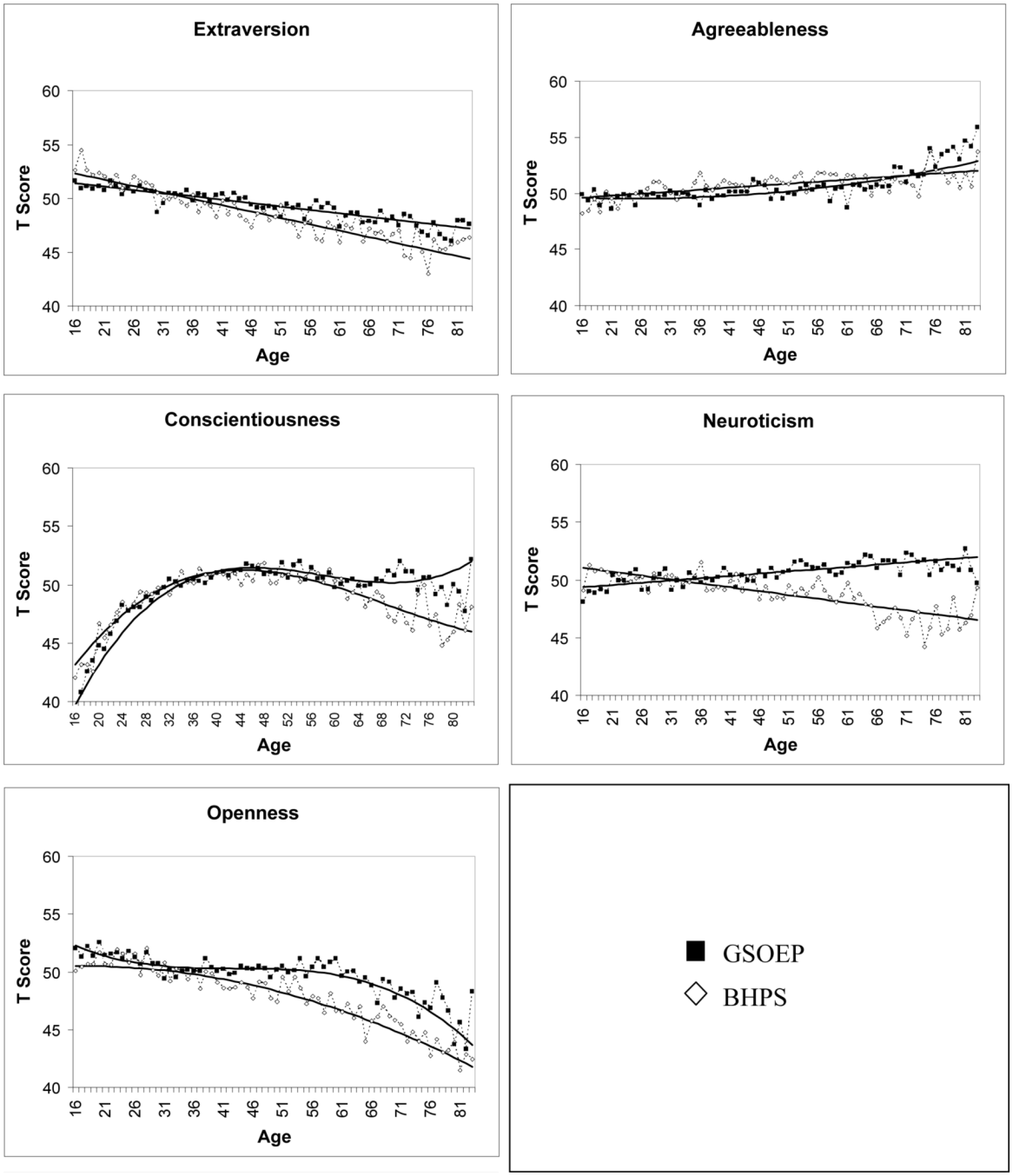

Figure 1.
Age Differences in the Big Five

Table 1

Means and Standard Deviations for the Big Five T Scores by Age Categories

	Minimum Sample Size	Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness
BHPS						
Age Group						
16−19	1,007	53.01 (9.86)	48.61 (10.61)	42.76 (10.97)	50.47 (10.86)	50.45 (10.43)
20−29	2,216	51.58 (9.93)	50.00 (10.11)	47.88 (10.11)	50.10 (10.36)	51.08 (10.21)
30−39	2,590	49.70 (10.33)	50.43 (10.02)	50.35 (10.21)	49.92 (10.03)	49.79 (10.16)
40−49	2,625	48.54 (10.94)	50.91 (10.03)	50.82 (10.62)	49.39 (10.34)	48.64 (10.56)
50−59	2,220	47.47 (11.03)	51.32 (10.51)	50.80 (10.99)	48.99 (10.81)	48.06 (11.45)
60−69	1,697	46.98 (11.29)	50.98 (11.37)	49.24 (12.32)	47.87 (11.17)	46.28 (12.54)
70−79	1,250	45.56 (12.35)	51.43 (11.96)	47.20 (13.21)	46.25 (11.22)	44.27 (13.58)
80−85	434	45.41 (12.10)	51.44 (11.92)	46.77 (13.29)	46.52 (11.74)	42.47 (12.86)
GSOEP						
Age Group						
16−19	1,344	51.17 (10.32)	49.64 (10.26)	41.49 (12.27)	48.80 (9.96)	51.75 (10.44)
20−29	2,835	50.94 (10.25)	49.65 (9.97)	47.15 (10.92)	49.99 (10.42)	51.46 (9.89)
30−39	3,745	50.12 (9.96)	49.79 (10.17)	50.22 (9.98)	50.04 (10.18)	50.23 (9.90)
40−49	4,275	49.84 (9.99)	50.31 (10.16)	51.22 (9.72)	50.36 (10.22)	50.15 (10.18)
50−59	3,271	49.08 (9.79)	50.21 (10.48)	51.16 (10.23)	51.10 (10.44)	50.43 (10.68)
60−69	3,293	48.27 (10.05)	50.56 (10.45)	50.23 (10.78)	51.51 (10.32)	49.43 (11.07)
70−79	1,683	47.54 (10.28)	52.46 (10.63)	50.46 (10.69)	51.38 (10.48)	47.66 (11.62)
80−84	403	47.57 (10.44)	54.16 (10.23)	49.84 (11.42)	50.74 (11.41)	45.56 (11.95)

Note: BHPS = British Household Panel Study; GSEOP = German Socio-Economic Panel Study; *T* scores were created by standardizing scores to the mean and SD for individuals ages 30 to 34 within each sample.

Psychol Aging. Author manuscript; available in PMC 2009 September 1.

Table 2

Regression Models Linking Age to the Big Five

		BHPS	GSOEP
Extraversion			
	Intercept	48.813 (.091)	49.526 (.070)
	Age	−.118 (.005)	−.063 (.004)
	R	.19	.11
	F for Linear	546.253	247.101
	ΔF for Quadratic	14.754	1.161
	ΔF for Cubic	2.267	0.128
Agreeableness			
	Intercept	50.655 (.089)	49.946 (.099)
	Age	.036 (.005)	.041 (.004)
	Age^2	-	.001 (.000)
	R	.06	.08
	F for Linear	54.214	109.273
	ΔF for Quadratic	12.481	36.543
	ΔF for Cubic	4.168	19.039
Conscientiousness			
	Intercept	51.252 (.133)	51.460 (.101)
	Age	−.019 (.012)	−.001 (.009)
	Age^2	−.007 (.000)	−.007 (.000)
	Age^3	.0001 (.000)	.0002 (.000)
	R	.20	.24
	F for Linear	31.280	449.358
	ΔF for Quadratic	483.736	596.991
	ΔF for Cubic	52.636	156.908
Neuroticism			
	Intercept	49.051 (.090)	50.541 (.071)
	Age	−.067(.005)	.039 (.004)
	R	.11	.07
	F for Linear	179.200	88.137
	ΔF for Quadratic	13.095	4.887
	ΔF for Cubic	0.441	2.360
Openness			
	Intercept	48.888 (.132)	50.284 (.101)
	Age	−.114(.005)	−.006 (.009)
	Age^2	−.002 (.000)	−.001 (.000)
	Age^3	-	−.0001 (.000)
	R	.20	.11
	F for Linear	525.857	199.951
	ΔF for Quadratic	29.079	15.202
	ΔF for Cubic	0.323	40.586

Note: Unstandardized coefficients reported. Standard errors reported inside parentheses. Age was mean-centered within each sample (Age in BHPS = 45.50; Age in GSOEP = 46.09). More complicated models involving age polynomials were selected only when the inclusion of the higher order term improved overall model fit at F > 25.00. Coefficients for age were reported from the final selected model within each sample. See text for complete details and justification of the model selection strategy.

Table 3

Effect Sizes for Gender and Education Differences by Age Categories

	Extraversion		Agreeableness		Conscientiousness		Neuroticism		Openness	
Gender Differences										
Age Group	*BHPS*	*GSOEP*	*BHPS*	*GSOEP*	*BHPS*	*GSOEP*	*BHPS*	*GSOEP*	*BHPS*	*GSOEP*
16−19	.38	.24	.29	.31	.13	.34	.57	.47	−.10	.36
20−29	.26	.13	.23	.30	.30	.13	.56	.47	−.18	.17
30−39	.27	.18	.39	.26	.16	.12	.45	.45	−.16	.12
40−49	.22	.23	.33	.37	.16	.13	.47	.35	−.21	.19
50−59	.16	.17	.38	.39	.08	.06	.64	.30	−.08	.09
60−69	.13	.14	.35	.42	.07	.09	.47	.37	−.14	.08
70−79	.08	−.01	.19	.38	−.08	.03	.51	.42	−.15	−.09
80−85/84	.06	.18	.22	.23	−.17	−.04	.39	.37	−.12	−.03
Overall	.20	.16	.31	.35	.11	.11	.51	.39	−.15	.12
Education Differences										
30−39	.20	.18	−.03	−.11	.17	.07	−.16	−.27	.51	.19
40−49	.27	.18	.04	.06	.34	.21	−.36	−.20	.62	.42
50−59	.11	.26	−.05	−.07	.14	.11	−.20	−.25	.52	.41
60−69	.10	.23	−.16	−.12	.05	.18	−.24	−.23	.54	.46
70−79	.07	.20	−.14	−.01	.07	.16	−.19	−.27	.37	.47
80−85/84	.03	.09	−.09	.08	.07	.22	−.20	−.23	.54	.30
Overall	.22	.16	−.10	−.06	.20	.22	−.10	−.17	.60	.32

Note: BHPS = British Household Panel Study; GSEOP = German Socio-Economic Panel Study. Effect sizes for gender were calculated so that positive scores indicated that women scored higher than men. Effect sizes for education were calculated so that positive scores indicated that more educated individuals scored higher than less educated individuals.