2008

# Ethnicity, Education, and the Temporal Stability of Personality Traits In the East Baltimore Epidemiologic Catchment Area Study

Corinna Löckenhoff, Antonio Terracciano, O. Joseph Bienvenu, Nicholas Patriciu, Gerald Nestadt, Robert R. McCrae, William W. Eaton, and Paul Costa

Published in final edited form as:
*J Res Pers*. 2008 ; 42(3): 577–598. doi:10.1016/j.jrp.2007.09.004.

# Ethnicity, Education, and the Temporal Stability of Personality Traits In the East Baltimore Epidemiologic Catchment Area Study

**Corinna E. Löckenhoff**[a,*], **Antonio Terracciano**[a], **O. Joseph Bienvenu**[b], **Nicholas S. Patriciu**[a], **Gerald Nestadt**[b], **Robert R. McCrae**[a], **William W. Eaton**[c], and **Paul T. Costa Jr.**[a,b,c]

[a]Laboratory of Personality and Cognition, National Institute on Aging, NIH, DHHS

[b]Department of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine

[c]Johns Hopkins University Bloomberg School of Public Health, Department of Mental Health

## Abstract

We examined the influence of age, gender, Black vs. White ethnicity, and education on five indices of personality stability and change across an average interval of 8 years in the East Baltimore Epidemiologic Catchment Area study. In the full sample (n = 505, aged 30-88), examination of structural, rank-order, ipsative, and mean level stability, as well as indices of reliable change suggested that NEO-PI-R personality traits showed moderate to high levels of stability over time. There were few age and gender effects on temporal stability but rank-order, ipsative, and mean level stability were lower among Blacks and individuals with lower education. Future research should explore additional demographic predictors of temporal plasticity in a diverse range of samples, and employ observer ratings to assess personality.

### Keywords

Personality; Five-Factor Model; Temporal Stability Indices; Demographics; Ethnicity; Education; Age; Epidemiologic Catchment Area

Questions regarding the extent of temporal stability of personality traits across the adult life course have important theoretical and practical implications. From a theoretical point of view, research on temporal stability and change and its predictors can shed light on the relative influence of maturational as compared to environmental factors on personality development (e.g., Ardelt, 2000). From a practical point of view, a better understanding of temporal stability and change in personality characteristics can inform therapeutic approaches to personality-related psychological problems and help to develop interventions tailored to the needs of individual clients (e.g., Widiger, Costa, & McCrae, 2002).

Given the importance of such questions, it is not surprising that recent years have seen a surge in meta-analyses examining longitudinal stability and change in personality traits (see Ardelt, 2000; Roberts & DelVecchio, 2000, Roberts, Walton, & Viechtbauer, 2006). However, the generality of the conclusions that can be drawn from previous studies is limited. First, the vast

*Corresponding author. Corinna E. Löckenhoff, Ph.D., Box #03, Gerontology Research Center, NIA, NIH, 5600 Nathan Shock Drive, Baltimore, MD 21224-6825, Phone: (410) 558 8196, Email: LoeckenhoffC@grc.nia.nih.gov.

majority of previous studies have relied on primarily White samples with high education levels (cf. Roberts et al., 2006) and apart from age and gender, few studies have rigorously examined the influence of potentially relevant demographic variables such as ethnicity and education on temporal stability of personality. Second, previous studies focused disproportionately on mean level and rank-order analyses of stability. Although such group level analyses arguably target the most fundamental aspects of personality development, other forms of stability or change, including individual level analyses, also merit attention (see De Fruyt, Bartels et al., 2006). Finally, most studies examined higher-level traits without consideration of lower order traits, or facets.

The present study was designed to address some of these gaps in prior research by examining the influence of four demographic variables (age, gender, education, and Black vs. White ethnicity) on five indices of plasticity/stability in a diverse community sample. To provide the rationale for our analyses, we discuss theoretical perspectives on the sources of personality stability and change and briefly review the literature on the association among demographic factors and different aspects of temporal stability of personality traits.

We adopted a trait approach to personality focusing on the Five-Factor taxonomy of personality, because it is a comprehensive and widely replicated model of personality structure (Paunonen, Zeidner, Engvik, Oosterveld, & Maliphant, 2000; McCrae et al., 2005). According to the Five-Factor Model (FFM; Goldberg, 1981, 1993; McCrae & Costa, 2003), personality can be described along five dimensions, Neuroticism (N), Extraversion (E), Openness to Experience (O), Agreeableness (A), and Conscientiousness (C). On a lower level of the hierarchy, each of these main dimensions is comprised of several individual facets. Developmentally, the adult personality system is thought to emerge from childhood temperamental differences and remain relatively stable during adulthood (e.g., McCrae & Costa, 2003).

When interpreting group-level findings concerning personality stability and change, researchers commonly distinguish between normative change, found to affect the majority of people in a similar fashion and at similar times in the life course, and non-normative change, concerning individual deviations from trajectories in the general population (e.g., Mroczek & Spiro, 2003). Recent theoretical discussions have focused on the relative contributions of biological as compared to environmental factors on temporal stability of personality (e.g., Ardelt, 2000; Roberts et al., 2006; Johnson, McGue, & Krueger, 2005) and both classes of explanations can account for normative as well as non-normative change. From a biological point of view, normative personality change is interpreted as the result of intrinsic trends based on rates of species-typical brain maturation and age-related changes in brain function and gene expression (e.g., McCrae et al., 2000), whereas non-normative change is viewed as the result of individual differences in genetics, injury, drug abuse, or disease processes that influence brain chemistry on an individual level (e.g., Dawson, Welsh-Bohmer, & Siegler, 1994; Piedmont, 2001). Environmental accounts, in contrast, interpret normative change as a consequence of universal patterns of life experiences and social roles (Elder, 1994), and non-normative change as the result of individual life experiences such as career trajectories (Roberts, 1997) or relationship patterns (Robins, Caspi, & Moffitt, 2002).

Although the relative contribution of biological and environmental factors to personality plasticity constitutes an important theoretical question, empirical investigations are often hampered by a wide-spread reliance on samples that are disproportionately wealthy and well-educated and underrepresent minorities (see Roberts et al., 2006, for a review). At present, the influence of sampling bias on conceptions of normative change is unknown because studies that examine the influence of demographic factors on indices of change have predominantly focused on age and gender -- only half of the studies on personality plasticity even report

information on ethnicity and education, let alone control for it (Roberts et al., 2006). This is an important omission since both biological and environmental sources of personality plasticity may differ considerably across demographic groups. For example, education levels and ethnicity are independently associated with exposure to environment toxins (e.g., Lanphear, Byrd, Auinger, & Schaffer, 1998) and use of psychoactive substances (National Institute on Drug Abuse, 2003). In addition to these biological factors, environmental factors such as career trajectories and social networks composition may also differ substantially across age, gender, ethnicity, and levels of education (Miech, Eaton, & Liang, 2003; Ajrouch, Antonucci, & Janevic, 2001a; 2001b).

In order to provide a comprehensive assessment of variations in temporal stability across demographic groups, the present study examined not only the commonly reported group-level indices of mean level and rank-order stability, but also measurement invariance as well as the individual-level indices of ipsative stability and reliable change. We now provide a quick overview of the different temporal stability indices and discuss their association with demographic factors.

One frequently neglected prerequisite for any analysis of longitudinal stability is measurement invariance (Horn & McArdle, 1992). Overall, estimates of the structural stability of FFM measures are encouraging and suggest that even when other types of change are found, there is little change in factorial structure across time (e.g., McCrae et al., 2002; De Fruyt, Van Leeuwen, Bagby, Rolland, & Rouillon, 2006). However, few previous studies have compared measurement invariance across demographic groups and the present study set out to address this important question.

Arguably the most thoroughly researched aspect of personality stability is rank-order consistency, the extent to which trait scores retain similar relative positions in a distribution of scores across two time periods. With regard to demographic predictors, most researchers agree that rank-order consistency (or test-retest stability) is considerably lower among adolescents and young adults than among older age groups (Ardelt, 2000; Roberts & DelVecchio, 2000). However, there is some disagreement regarding the age(s) or life-stage when the age-related increase in the magnitude of the test-retest stability coefficients peaks. Some studies have reported that stability continues to increase from early adulthood to age 50 (Roberts & Del Vecchio, 2000), some have found a plateau, with the stability coefficients reaching their maximum magnitude after age 30 (Terracciano, McCrae, & Costa, 2006), and others have reported curvilinear trajectories with increases until age 50 but declines in old age (Ardelt, 2000). Studies have found similar patterns of rank-order stability for men and women (Roberts & DelVecchio, 2000), but the influence of ethnicity and education has not been examined. Given that other individual difference variables such as global well-being tend to show somewhat lower rank-order stability among Blacks as compared to Whites and among participants with lower education levels (e.g., Costa, McCrae, & Zonderman, 1987), this remains an important open question.

Another highly researched aspect of temporal stability at the group level is change in mean levels of traits. With regard to demographic characteristics, cross-sectional (McCrae & Costa, 2003; McCrae et al., 1999; Srivastava, John, Gosling, & Potter, 2003) and longitudinal studies (Terracciano, McCrae, Brant, & Costa, 2005; Roberts et al., 2006) suggest that the greatest amount of change happens in young adulthood, but there continue to be modest declines in N, O, and some facets of E, as well as increases in A and C in middle and late adulthood. On average, the magnitude of such changes tends to be small (Costa, Herbst, McCrae, & Siegler, 2000; Roberts et al., 2006) although the specific patterns of change vary somewhat across studies. Further, some studies show modest gender differences in mean level changes. Most notably, women were found to show age-related decreases in N and emotional dependence,

whereas men remained stable in those traits (Viken, Rose, Kaprio, & Koskenvuo, 1994; Wink & Helson, 1993; but see Roberts et al., 2006). Although the influence of age and gender on mean-level stability is relatively well-researched, little is known about education levels and ethnicity as moderators of mean level stability.

Finally, it is possible to examine stability and change at an individual level. Using the Reliable Change Index (RCI; Jacobson & Truax, 1991) to identify whether a given individual shows a level of change above chance fluctuations, previous studies have found that although more individuals show reliable change than would be expected by chance, the majority of participants remain stable (Roberts, Caspi, & Moffit, 2001; Robins et al., 2001; Vaidya, Gray, Haig, & Watson, 2002; DeFruyt, Bartels et al., 2006). In previous research, reliable change in personality characteristics was shown to be sensitive to therapeutic interventions (De Fruyt, Van Leeuwen et al., 2006), but again, little is known about its association with demographic factors. Including this individual-level index of temporal stability in the present analyses is important because it will allow us to determine whether any differences across demographic groups are disproportionately due to a few select individuals or whether they are seen consistently among larger proportions of the sample.

In addition to examining reliable changes in single traits, there are methods that target the complete profile of an individual's traits. This is important because the implications of a person's standing on a given trait may, at least in part, depend on his or her standing on other traits (Costa & Piedmont, 2003). Indices of ipsative stability, that is consistency in the configurations of traits within individual participants, can address this concern. Ipsative stability is usually measured through intraclass correlations (*ICC*) computed for each individual across the whole range of her or his traits measured at two time points. The finding that *ICC*s are higher among college students followed over a four-year interval (Robins, Fraley, Roberts, & Trzesniewski, 2001, *Mdn* = .76) than among depressive patients assessed before and after treatment (De Fruyt, Van Leeuwen et al., 2006, *Mdn* = .49) suggests that ipsative stability may help in capturing individual differences in profile consistency. However, to the best of our knowledge, temporal variations in this index of individual stability or its relation to demographic variables has not been extensively examined.

The overarching goal of the present study was to examine different aspects or indices of temporal stability of personality traits at the group and individual level in a diverse sample of participants and to examine whether stability indices differed systematically across demographic groups, with particular emphasis on the role of education and ethnicity. Across an average time interval of eight years, we followed an adult community sample representative of a neighborhood in inner city Baltimore. Participants were of mixed ethnic descent, and spanned a wide range of education levels.

Because previous studies have found differences in personality stability across individual facets of the FFM (e.g., Terracciano, McCrae, Brant, & Costa, 2005) we examine stability indices at both the factor and the facet level. Further, as discussed above, examining only mean level changes and rank-order stability does not paint a complete picture of temporal stability or change. Therefore, we also examined structural stability, reliable change, and ipsative stability (cf. De Fruyt, Bartels et al., 2006). It is, however, important to note that the present study was limited to two data points and could therefore not examine the shapes of trajectories over time.

As discussed above, environmental and biological sources of personality change differ across demographic groups and we therefore expected to find subtle differences in stability indices by age, gender, Black vs. White ethnicity, and education. In general, we expected to replicate previous findings regarding the influence of age and gender on temporal stability as described above. Although there is limited research on the influence of ethnicity and education on

temporal stability indices, we expected to find somewhat lower stability rates and less adaptive trajectories among participants with lower education and among Blacks, because limited access to financial resources (in the case of participants with low education) and societal stereotypes (in the case of Blacks; e.g., Major & O'Brien, 2005) leave these groups with fewer resources to buffer the influence of environmental and biological sources of non-normative change. Nevertheless, given that heritability for personality characteristics is high (Loehlin, McCrae, Costa, & John, 1998) and similar maturational patterns are found across a diverse range of societies (McCrae & Costa, 2006; McCrae et al., 1999), we predicted that demographic variables would only account for a small proportion of the variance in temporal stability.

## Method

### Participants and Procedure

The present study draws its sample from the East Baltimore Epidemiologic Catchment Area study (ECA, Eaton, et al., 1997), an ongoing multidisciplinary study of community-dwelling residents of East Baltimore. A probability sample of 3,481 household residents was initially interviewed in 1981 and followed up in 1993-98 ($n$ = 1,920) and in 2004-2005 ($n$ = 1,071).

Personality assessments were available for a sub-sample of 573 participants during the two follow-ups.[1] The present study is limited to participants who had valid personality assessments and Mini Mental State Scores (Folstein, Folstein, & McHugh, 1975) above the cutoff value of 23 at both times of assessment. Also, to allow for meaningful comparison among Blacks and Whites, we dropped the participants who did not belong to either of these ethnic groups or whose ethnicity was unknown (3% of otherwise elegible participants). The final sample size for the present study was n = 505.

The present sample was composed of 64% Whites and 36% Blacks. Sixty-seven percent were women. Participants' age at the first personality assessment ranged from 31-88 years ($M$ = 46.65, $SD$ = 1.57). The mean retest interval was 8.4 years ($SD$ = 2.01).[2] Education ranged from 5 to 17 years ($M$ = 12.90, $SD$ = 2.39); 22% of participants had less than 12 years of schooling, 34% had a high school degree, 37% had at least some college, and 7% had a postgraduate education. For some of the analyses, age was dichotomized at age 45 and the education variable was split at the high school level. Table 1 (left columns) shows sample sizes and percentages for each of the demographic groups.

Participants who only completed the initial wave of the Baltimore ECA and did not participate in follow-up assessments were significantly older, $t(3,479) = 18.60$, $p < .001$, $d = .64$, and less educated, $t(3,478) = 14.50$, $p < .001$, $d = -.50$, than participants who remained in the study, but there were no statistically significant differences in gender or ethnic composition (compare also Badawi, Eaton, Myllyluoma, Weimer, & Gallo, 1996).

[1]The most recent personality assessment was conducted as part of the main Baltimore ECA interview. Thus, all participants completed the questionnaire under identical conditions. The initial personality assessment, in contrast, was conducted in the context of several different projects. Fifty percent of the participants in the present sample completed the personality measure during a project on Axis I psychological disorders (Bienvenu et al, 2001), 31% were assessed as part of a project on the prevalence and correlates of personality disorders in a community sample (Samuels et al. 2002), and 19% were assessed in the context of a project on marijuana use (Ford, Vu, & Anthony, 2002). Because personality traits did not differ significantly across the different projects, data were collapsed for further analyses.

[2]As indicated in Footnote 1, data assessment at Time 1 was conducted in the context of several different projects. As a result, the distribution of retest intervals was bimodal. After dichotomizing the variable at the mean, we conducted preliminary analyses to examine associations with the predictor and outcome variables. Age, education, gender, ethnicity, and personality scores did not differ significantly by retest interval. Furthermore, the retest interval did not predict rank order stability scores, ipsative stability scores, the degree of mean level change, or the number of facets showing reliable change. For the sake of simplicity we therefore did not include this variable in further analyses.

Among the participants who completed the two follow-up assessments, those who were not administered the NEO-PI-R or did not meet the criteria for inclusion in the present sample for other reasons were significantly older, $t(1,069) = 8.90$, $p < .001$, $d = .54$, and less educated, $t(1,069) = 5.60$, $p < .001$, $d = -.34$. Again, there were no gender differences; and, after accounting for the fact that we had selectively eliminated participants who were neither Black nor White, the two groups did not differ significantly by ethnicity.

### Measures

**Personality—**Personality was assessed with the Revised NEO Personality Inventory (NEO-PI-R, Costa & McCrae, 1992), a 240-item measure of the Five-Factor Model of personality (FFM; McCrae & John, 1992). It contains 30 eight-item facet scales, six for each of the five basic personality factors: N, E, O, A, and C. Domain scores are the sum of the six relevant facet scores; more precise factor scores are calculated from factor scoring weights given in the manual. Scores for the factors and facets were standardized as *T*-scores ($M = 50$, $SD = 10$) using adult combined-sex norms reported in the manual (Costa & McCrae, 1992).

Participants completed the questionnaires in a quiet setting, either at their own home or at a participating institution. NEO-PI-R items are answered on a five-point Likert scale, from *strongly disagree* to *strongly agree*, and scales are roughly balanced to control for the effect of acquiescence. Although the NEO-PI-R does not include a social desirability scale (Piedmont, McCrae, Riemann, & Angleitner, 2000), it does provide checks for protocol validity based on the number of missing items, direct questions regarding honesty and accuracy of the response, and examination of random responding patterns.

The NEO-PI-R has been translated into multiple languages and used in more than 50 cultures (McCrae, et al., 2005). Evidence on convergent and discriminant validity is presented in the Manual (Costa & McCrae, 1992).

### Analyses

Our goal of examining the influence of four demographic factors on a comprehensive set of personality stability indicators – both on the factor and the facet level – required multiple comparisons for each indicator. Given the scarcity of previous research regarding the influence of ethnicity and education on personality stability and the exploratory nature of our analyses we opted to minimize the probability of Type 2 errors (i.e., erroneously assuming that personality has remained stable) by presenting the findings without controlling for familywise Type 1 error rate (i.e., Bonferroni correction) or false discovery rate (Benjamini & Hochberg, 1995). Instead, we provide detailed information about effect sizes to put our findings into perspective. This allows future researchers to develop targeted hypotheses which can be tested using more stringent criteria.

## Results

### Data Quality and Structural Stability

Because the present sample was more diverse than most previously studied samples (cf. Roberts et al., 2006) and because participants were comparatively low in education, our first goal was to establish that data quality was sufficient to conduct the analyses we had planned. To this end, we examined the percentage of valid protocols, acquiescence, internal consistency, congruence with the normative factor structure, and structural stability over time. Due to a clerical error, raw data were partially lost for the Time 1 assessment (i.e., the first follow-up). Thus, some aspects of data quality could only be examined for Time 2 (i.e., the second follow-up).

Overall, 96% of the participants who completed the NEO-PI-R at Time 2 had valid protocols, and 4% showed evidence of acquiescence. Acquiescence and proportions of valid questionnaires at Time 2 did not differ by age group, gender, or ethnicity; but compared to college-educated participants, participants at high-school level or below had higher acquiescence rates (5% vs. 2%, $\chi^2 = 3.92$, $p < .05$) and lower proportions of valid questionnaires (95% vs. 98%, $\chi^2 = 4.60$, $p < .05$).

Internal consistencies for the five domain scores at Time 2 were satisfactory as well. For the most recent NEO-PI-R assessment in 2004/05, Cronbach's α for the domains ranged from .84 (for O) to .92 (for N). Table 1 shows mean personality and internal consistency scores by demographic group at Time 2. There were few differences by age, gender, and education. Blacks had slightly lower coefficients than Whites, especially on O and A, and an examination of the standard deviations in Table 1 suggests that this could be due to somewhat lower variability among Blacks. Nevertheless, all of the internal consistency scores were .79 or higher which is close to the scores reported in the NEO-PI-R Manual (Costa & McCrae, 1992) indicating that internal consistency was adequate for the sample as a whole as well as for the demographic subgroups.

Although the factor structure of the NEO-PI-R has been replicated across different age, gender, cultural, and clinical groups, the demographic characteristics of the Baltimore ECA sample and the fact that personality was assessed at two time points represent an interesting test of factorial invariance. To examine factorial invariance we rotated the factor structures at each time of assessment towards the factor structure reported for the normative sample (Costa & McCrae, 1992) for maximum alignment (McCrae, Zonderman, Costa, Bond, & Paunonen, 1996). At Time 1, factor congruence coefficients were acceptable for all groups (all coefficients ≥ .90). At Time 2, Blacks showed congruence coefficients for E (.83) and O (.84) that were just below the .85 criterion suggested by Haven and ten Berge (1977). All other coefficients were ≥ .92. These coefficients are comparable to those found in another African-American sample (Salva, Davey, Costa & Whitfield, 2007).

To examine structural stability over time, we rotated the factor structure at the follow-up assessment (2004/05) toward the structure for the initial assessment (early 1990s) for maximum alignment (see McCrae et al., 1996 for a detailed description of the statistical procedures). In the sample as a whole, factor congruence coefficients were .98, .98, .98, .99, and .98 for N, E, O, A, and C, respectively. Facet congruence coefficients, which capture the congruence of each facet's loadings across the five factors with the corresponding loadings in the normative matrix, were .97 or higher. These high coefficients support the factorial invariance of the NEO-PI-R in the present data and present evidence of structural invariance across time. We also examined structural stability within each of the demographic groups. Among Blacks, structural stability was .86 for E and .78 for O. This is consistent with the slightly lower congruence coefficients for these traits among Blacks at Time 2. All other structural stability coefficients were ≥ .94.

In summary, the data quality index, internal consistencies, congruence coefficients, and indices of structural stability painted a consistent picture suggesting that data quality in the present sample was adequate to conduct the other stability analyses we had planned.

## Rank Order Stability of Factors and Facets

Rank order stability coefficients were computed to examine changes in participants' relative standing within the group across the 8 year interval. These are presented in the second column of Table 2 for the total sample for each of the factors and facets. All of the factor coefficients were above .63 and all facet coefficients were above .50, suggesting moderate levels of stability over time.

It would be possible to compare stability across groups by dichotomizing demographic variables and computing retest coefficients within group. We did so in preliminary analyses, which showed that retest stability was generally lower in Blacks (Mdn = .52) than in Whites (Mdn = .64), and lower in less educated (Mdn = .57) than in more educated respondents (Mdn = .65). However, the demographic variables are themselves correlated, and dichotomizing continuous variables such as age and education fails to utilize all the information contained in the total distribution of scores. We chose instead to quantify stability at the level of the individual, and to predict individual differences in stability from the set of demographic variables.[3]

We computed individual stability scores for each participant on each of the NEO-PI-R factors and facets. For this purpose we adapted the Ipa statistic (McCrae, 1993) for use with a single trait score assessed at two time points (see Costa & Piedmont, 2003 for a precedent to this approach):

$$I_{pa}=\frac{1+2M^2-d^2}{\sqrt{10}},$$

where $M$ is the mean of the $z$-standardized personality trait scores for a given individual at the two times of administration and $d$ is the difference between those scores.[4] Mathematically, the mean of the $I_{pa}$ measure across individuals is proportional to the correlation $r_{12}$ between Time 1 and Time 2 scores which facilitates interpretation:

$$\overline{I_{pa}}=\frac{3}{\sqrt{10}}r_{12}\approx 0.95r_{12}$$

We computed regression analyses with individual stability scores for each of the factors and facets as the dependent variables. Age, gender, ethnicity, and years of education were simultaneously entered as predictors. The last five columns of Table 2 provide a summary of the results.

Age was associated with higher individual stability in the E factor, but no other effects of age were found. Gender differences were limited to two facets: Women showed lower stability than men in A2: Straightforwardness and higher stability in C2: Order. Compared to Blacks, Whites had higher individual stability scores on three of the factors: N ($M$ = .71 vs. $M$ = .52), E ($M$ = .77 vs. $M$ = .46), and A ($M$ = .79 vs. $M$ = .47). On the facet level, Whites were more stable on N1: Anxiety, three facets of E, two facets of O, two facets of A, and C2: Order. Higher education levels were associated with greater stability for three of the factors; compared to participants educated at high school level or below, college educated participants had higher individual stability scores on E ($M$ = .76 vs. $M$= .59), O ($M$ = .85 vs. $M$ = .60), and C,($M$ = .69

[3]We also considered a moderated regression approach and preliminary analyses indicated that the general pattern of results was comparable to findings obtained from $I_{pa}$ scores. However, other than moderated regressions, individual stability scores allowed us to quantify each participant's contribution to overall stability at an individual level and directly relate it to individual differences in demographic variables. Also, the regression approach implicitly assumes that inter-individual differences in stability are random whereas the theoretical considerations outlined in the introduction suggest that demographic differences in stability may reflect meaningful changes in personality traits.

[4]Compared to other measures of individual stability, such as Asendorpf's (1992) measure, $I_{pa}$ has the advantage that it is sensitive not only to absolute differences between $z$-scores, but also to the extremeness of scores. Thus, it accounts for the fact that an individual who has $T$-scores of 80 and 81 on two occasions offers more striking evidence of stability and contributes more to the retest correlation than an individual who scores 50 and 51. Asendorpf's measure does not.

vs. $M = .53$). College educated participants also scored higher on two facets of E, five facets of O, and on C2: Order.

It is important to note that although demographic variables were significant predictors of rank-order stability in personality scores, they did not predict more than 5% of the variance in the stability of any of the factors or facets.

## Ipsative Stability of Personality Profiles

We next examined profile consistency as an index of ipsative stability. For this purpose, we computed ICCs across the 30 facets of the NEO-PI-R (see McCrae, in press; Terracciano & McCrae, 2006 for a detailed discussion of the merits of ICC over other measures of profile consistency). ICC scores for each participant were derived by computing a simple Pearson correlation across the 30 NEO-PI-R *T*-score standardized facets with the double entry method. The ICCs for the 505 individuals ranged from −.42 to .94 ($Mdn = .60$, $M = .53$; $SD = .25$) and only 4% of the sample showed negative stability coefficients.

To examine if ICC was associated with any of the demographic characteristics, we computed a regression with ICC as the dependent variable and age, gender, ethnicity, and education as predictors. To account for negative skewness, ICC scores were first reflected by subtracting them from a constant k equaling the largest score +1 (in this case k = 2) and then log transformed. Results suggested that ICC was positively associated with age ($\beta = .09$, $p < .05$), years of education ($\beta = .16$, $p < .001$), and White ethnicity ($\beta = .22$, $p < .001$). There were no gender differences in profile consistency. Overall, demographic variables accounted for 9% of variability in ICC scores ($R^2 = .09$, $p < .01$).

## Mean-Level Stability of Factors and Facets

For the sample as a whole, Table 3 shows mean scores and standard deviations for the personality factors and facets at both times of assessment. Paired samples *t*-tests in which participants served as their own controls indicated that over time, there was a 1.2 *T*-score point increase in N, a .6 point decrease in O, and a .8 point increase in C. On the facet level, there were significant increases in four facets of N (N1: Anxiety, N2: Angry Hostility, N3: Depression, and N6: Vulnerability), significant decreases in three facets of E (E4: Activity, E5: Excitement Seeking, and E6: Positive Emotions) and O4: Actions, and significant increases in A5: Modesty and C3: Dutifulness. As at the factor level, the magnitude of these changes was quite small. The largest change was seen for N1: Anxiety (2.1 *T*-score points) followed by O4: Actions (1.4 *T*-score points), E4: Activity (1.1 *T*-score points), and C3: Dutifulness (1.0 *T*-score point). All other changes were less than one *T*-score point.

Whereas the slight decline in O and the increase in C are consistent with the results of previous studies, the increase in N and four of its facets is an atypical finding (cf. Roberts et al., 2006). To examine whether the observed mean-level changes were more pronounced or disparate in certain demographic groups, we computed ANOVAs with time of assessment as a repeated measures variable, the demographic variables as between-subject factors, and the personality scores for each of the factors and facets as the dependent variables. The models included the interactions of assessment time with demographic factors but none of the higher order interactions among demographic variables. Table 4 reports effect sizes and significance levels for each of the four Time × Demographic interactions. The main effects of time are not shown since they are identical to the ones presented in Table 3.

For age, there was an interaction for O2 suggesting that older adults showed a decline in Openness to Aesthetics while younger adults remained stable. For gender, there were interactions for N3 and N6, suggesting that women increased in Depression and Vulnerability

while men remained stable, and for O3 suggesting that women remained stable in Openness to Feelings while men decreased. For ethnicity, there was an interaction for the N factor and two of its facets (N1: Anxiety and N3: Depression) suggesting that Whites increased in N while Blacks decreased. Interactions for O6 and A1 suggested that Whites decreased in Openness to Values and Trust, while Blacks showed a slight increase in these facets. Blacks also increased in two facets of C while Whites remained stable. For education, there were interactions for the O factor, O1: Fantasy, and C2: Order suggesting that more highly educated participants showed a decrease in these traits whereas participants with lower education showed an increase.

To summarize, these findings suggest that mean levels of personality were generally stable over time. Further, mean level change was significantly associated with demographic factors, most notably ethnicity, but the corresponding effect sizes were very small (all partial $\eta^2 < .02$, see Table 4) accounting for less than 2% of the variance.

## Reliable Change in Individual Scores

As a final examination of temporal consistency we examined the Reliable Change Index (RCI; Jacobson & Truax, 1991), to identify the proportions of individuals whose personality changes were larger than fluctuations expected by chance. Reliable change scores were computed by dividing the difference between pretest and posttest by $S_{Diff}$, the standard error of the difference between the test scores.[5]

$$RCI = \frac{X_2 - X_1}{S_{Diff}}$$

Consistent with the literature (Jacobson & Truax, 1991; Roberts et al., 2001; Robins et al., 2001; Vaidya et al., 2002; De Fruyt, Bartels et al., 2006), a criterion of 5% for establishing reliable change was adopted. Thus, participants with RCI scores larger than 1.96 or smaller than −1.96 were classified as showing reliable change.

For each of the factors and facets, Table 5 reports the percentages of participants that showed reliable decrease, remained stable, or showed reliable increase. Although the majority of participants remained remarkably stable, all factors and facets, with the exception of A2: Straightforwardness, showed larger proportions of changing individuals than would have been expected by chance. In general, the proportion of participants showing increase as compared to decrease paralleled the observed mean level changes. For example, consistent with the mean level increase in N, 1.5 times as many participants showed increase as compared to decrease in N.

---

[5]Following Jacobson and Truax (1991), $S_{Diff}$ was derived from the following formulas:

$$S_{Diff} = \sqrt{(2 * SE^2)}$$

$$SE = SD\sqrt{(1 - r_{tt})}$$

We used 2-year retest correlations (from McCrae, Yik, Trapnell, Bond, & Paulhus, 1998) to estimate $r_{tt}$, and, since we were using *T*-scores, the *SD* was set to 10. Note that the $r_{tt}$ scores reported by McCrae et al. are based on a sample of Baltimore Longitudinal Study of Aging participants who are disproportionately well-educated and may therefore show higher data quality and retest-reliability than the present sample. Thus, although the present RCI scores represent the best available estimate, they may overestimate actual change.

To examine reliable change across all NEO-PI-R facets, we computed a summary score indicating on how many of the 30 facets each participant had shown reliable change (cf. De Fruyt, Van Leeuwen et al., 2006). On average, participants changed on fewer than 4 of the facets ($M = 3.58$, $SD = 3.34$). To examine the influence of demographic characteristics on reliable change, we computed a regression with the summary score as the dependent variable and age, gender, ethnicity, and education as predictors. Since the summary variable was slightly skewed to the left, log transformed values were used in this analysis. Results suggested that the number of facets showing reliable change was negatively associated with age ($\beta_{Age} = -.10$, $p < .05$), and years of education ($\beta_{Education} = -.11$, $p < .05$). In other words, older and more highly educated participants included fewer reliable changers. Ethnicity and gender were unrelated to the summary score and overall, demographic variables accounted for only 4% of the variability in the summary score ($R^2 = .04$, $p < .01$).

## Discussion

The present study examined the influence of demographic variables on different aspects of temporal stability and change in personality traits in a community sample of East Baltimore residents followed across an 8-year time interval. It extends previous work by systematically examining not only age and gender but also the influence of education and Black vs. White ethnicity on personality stability or change. Our findings also add to previous work by examining not only mean-level and rank order stability but also three additional approaches to assessing stability and change at the factor and the facet level. After briefly summarizing our findings for the sample as a whole, we consider differences in stability indices across demographic groups and discuss limitations as well as implications for future research.

### Five Aspects of Personality Plasticity in the Baltimore ECA

Overall, participants in the Baltimore ECA sample remained relatively stable. In the sample as a whole, structural stability coefficients were excellent, ranging from .98 to .99. Rank order stability on the factor level ranged from .64 to .75. As expected, these coefficients are somewhat lower than those reported for less diverse and more educated samples (e.g., Costa et al., 2000). Nevertheless they are well within the typical ranges reported in the literature (Roberts & Del Vecchio, 2000; Ardelt, 2000).

When comparing personality profiles over time, we found a median ICC of .6. Again, this value is somewhat lower than the median of .76 reported for ipsative stability of the Big Five in student samples (Robins et al., 2001), but well above an ICC of 0.0 which would be expected if the *T*-standardized profiles were paired up randomly.

The magnitude of mean level changes for the factors ranged from .01 *T*-score points for A to 1.2 *T*-score points for N. This is consistent with previous research suggesting that mean level changes are small in magnitude (Roberts, et al. 2006; Costa et al., 2000). Also consistent with the previous literature, we found increases in C, and decreases in O and three facets of E. However, in contrast to previous studies, we found an increase in N and four of its facets, but no increases in A.

For each of the factors and facets, reliable change was found for less than one in five participants. On average, participants changed on fewer than four of the facets. These percentages are comparable to the values reported by Vaidya et al. (2002) in a student sample followed over 2.5 years and suggest that although the percentages of individuals who showed reliable changes was higher than would have been expected by chance, the majority of the sample remained stable.

Taken together, these results support the notion that personality remained relatively stable in this urban sample. This is an important finding because it suggests that previous studies that found high levels of stability in mostly college educated samples appear to generalize to a more diverse community sample.

## Demographic Characteristics and Personality Plasticity

Since the main goal of the present study was to compare indices of temporal stability and change in FFM assessed personality traits across demographic groups, especially with regard to education and Black vs. White ethnicity, we now provide a brief summary of results for each demographic factor to illustrate similarities and differences across various indices and relate our findings to the previous literature.

As found previously (Roberts & Del Vecchio, 2000), increasing age was associated with greater rank-order stability, although in the present study, this effect only reached statistical significance for a single factor: E. Also, older adults showed greater ipsative stability, and lower frequencies of reliable change. With regard to mean level changes, the older age group showed decreases in O2: Openness to Aesthetic while the younger age group remained stable. No other age effects were found. Thus, while age effects in the present study generally conformed to previously found trajectories, these effects were relatively few in number and small in size.

Consistent with prior research (Roberts & Del Vecchio, 2000) gender was not related to rank-order stability on the factor level. Similarly, on the facet level, only 2 out of 30 possible facets showed significant gender differences. With regard to mean level stability, we did not replicate prior findings suggesting that women decline in N while men remain stable (e.g., Helson & Wink, 1992). Instead, there were no associations between gender and mean level stability on the N factor, and on the facet level, women actually increased in N3: Depression and N6: Vulnerability, whereas men remained stable. Further, women remained stable in Openness to Feelings (O3) while men declined. Gender was unrelated to ipsative stability and reliable change. Thus, gender differences in the present study were comparable to previous studies with regard to rank-order consistency, but showed divergent trajectories for mean level stability for three of the facets.

When comparing stability indices across ethnicity, we found that Whites were more stable than Blacks in rank-order and ipsative stability. Reliable change, in contrast, was unrelated to ethnicity suggesting that the observed differences in stability are not due to select individuals but reflect general trends. Mean level trajectories for Whites and Blacks differed for the N factor and two of its facets, suggesting that Whites showed a slight increase in N whereas Blacks decreased. Mean level changes also differed for two facets of C, with Blacks showing increases while Whites remained stable. Thus, while Blacks were generally less stable than Whites, they showed greater consistency with previously reported mean level changes, especially for N and C.

Finally, temporal stability of personality factors and facets varied across educational levels. Higher education was associated with greater rank-order stability, ipsative stability, and lower frequencies of reliable change. Further, while more highly educated participants decreased in O, O1: Openness to Fantasy, and C2: Order, participants with lower education levels showed the opposite trajectories.

The general finding that more temporal instability was seen in Black and lower education groups may have two explanations. An *artifactual* explanation would argue that individuals in these groups have less test-taking experience or less motivation to cooperate fully, and as a result, produced lower quality data and therefore less consistent results. This argument is most

compelling for the low education group, where somewhat higher acquiescence and a lower proportion of valid protocols were seen. Further, although ethnic groups did not differ in acquiescence or the proportion of valid questionnaires, Blacks showed somewhat lower congruence coefficients and structural stability on E and O, and lower internal consistencies and levels of variability on all of the factors. Such problems could be circumvented in future studies by obtaining observer ratings from individuals who are trained to produce high quality data.

A *substantive* explanation for demographic differences in plasticity would suggest that life circumstances are different, and perhaps more challenging, for certain demographic groups. As discussed in the introduction, both biological and environmental sources of non-normative personality plasticity may differ considerably across demographic groups and future studies are needed to examine the influence of such variables. The role of biological factors, for example, could be explored by accounting for exposure to psychoactive substances and environmental toxins. With regard to environmental factors, future studies could examine the experience of stressful life events and the available resources that could help to buffer such events (e.g., financial assets, social support). In this context, it is particularly intriguing that mean level trajectories for Blacks showed greater similarity to the patterns of adaptive changes (i.e., increases in C, and decreases in N) that were observed in previous studies with predominantly White samples, whereas Whites in the present study showed a potentially maladaptive increase in N. The greater resilience observed among Blacks in this predominantly urban sample may be associated with culture-specific coping patterns (Utsey, Bolden, Lanier, & Williams, 2007) --another potentially fruitful avenue for future research.

In summary, these findings support our prediction that indices of temporal stability and change would differ not only by age and gender but also across ethnic groups and levels of education. Nevertheless, it is important to keep in mind that across the different indices, demographic factors only accounted for a very small portion of the variability in personality plasticity (i.e., $< 2\%$ for mean level plasticity, $\leq 5\%$ for rank order plasticity and reliable change, and $< 10\%$ for ipsative stability).

## Limitations and Future Directions

It is important to acknowledge several limitations of the present study. Most importantly, we examined only two data points. Thus, we could not determine whether specific demographic groups showed similar trajectories that merely differed in the age of onset and acceleration, or whether there were differences in the actual shapes of the trajectories. Because recent hierarchical linear modeling analyses suggest that N, E, and C follow curvilinear trajectories in adulthood (Terracciano, et al., 2005) this constitutes an important direction for future research.

Another concern associated with the examination of a single time period is that results may have been influenced by community and societal events occurring between the two times of assessment. The observed increase in N, for example, could be associated with the events of 9/11/2001 as well as the beginning of the Iraq war which both occurred within 5 years of the second wave of data collection. This interpretation is supported by the finding that the increase in N was particularly pronounced for the facet N1: Anxiety. More subtle societal changes such as a loss in consumer confidence in the aftermath of the “dot.com” crash and increasing unemployment rates due to a decline in the local industry may have influenced our findings as well. Future studies need to examine multiple time points to disentangle the influence of societal events from other sources of personality plasticity.

Also, although the present study is the first to report separate indices of personality continuity across education levels and ethnic groups in a community sample, our sample is not

representative of the general U.S. population. Moreover, in spite of their ethnic and educational diversity, participants were drawn from a single area in downtown Baltimore and therefore from a relatively homogenous cultural background which could have obscured the role of environmental effects. Future research should strive to examine other important minority groups such as Asian and Hispanic Americans and draw participants from across the U.S.

Our findings are also potentially limited by our use of the NEO-PI-R to assess personality. For one, there are alternative personality questionnaires (e.g., the MPQ, Tellegen, 1982, or the EPQ, Eysenck & Eysenck, 1974) that propose a different structure of personality. However, since the factors and facets of the NEO-PI-R are quite inclusive and the FFM maps well to the content domains of other measures (e.g., John, 1990) our findings can inform research in the context of other personality models as well. Further, the present study relied on a self-report measure of personality and the observed effects may partially reflect changes in self-views as opposed to actual changes in personality. Future studies could address this issue by using observer-ratings as well as self-reports.

In spite of these limitations, the present findings contribute significantly to our understanding of personality stability or plasticity over time and suggest promising avenues for future research. Perhaps the most interesting finding was that among the different demographic characteristics, education and ethnicity showed stronger associations with rank-order, ipsative, and mean level stability than age and gender. Given the large body of literature that has already been devoted to relatively subtle differences in personality plasticity across age and gender, it is important to conduct further research to replicate the present findings and explore additional demographic predictors of personality plasticity (e.g., income, job status) in a diverse range of samples. However, as discussed above, the magnitude of the observed effects was very small, and it appears that similarities in personality trajectories across demographic groups outweigh any differences.

## Acknowledgments

Robert R. McCrae and Paul T. Costa, Jr., receive royalties from the Revised NEO Personality Inventory. This paper was supported by National Institutes of Health grants MH 47447, MH64543, and MH 50616 and by the Intramural Research Program of the NIH, National Institute on Aging. We thank Angelina Sutin and Matthew Scally for their comments on earlier versions of this draft.

## References

Ajrouch KJ, Antonucci TC, Janivic MR. Social networks among blacks and whites: The interaction between race and age. Journals of Gerontology: Social Sciences 2001a;56B:S112–S118.

Ajrouch KJ, Antonucci TC, Janivic MR. Social networks among men and women: The effects of age and socioeconomic status. Journals of Gerontology: Social Sciences 2001b;60B:S311–S317.

Ardelt M. Still stable after all these years? Personality stability theory revisited. Social Psychology Quarterly 2000;63:392–405.

Asendorpf JB. Beyond stability: Predicting inter-individual differences in intra-individual change. European Journal of Personality 1992;6:103–117.

Badawi MA, Eaton WW, Myllyluoma J, Weimer LG, Gallo J. Psychopathology and attrition in the Baltimore ECA 15-year follow-up 1981-1996. Social Psychiatry & Psychiatric Epidemiology 1999;34:91–98. [PubMed: 10189815]

Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society 1995;57:289–300.

Bienvenu OJ, Nestadt G, Samuels JF, Howard WT, Costa PT Jr, Eaton WW. Phobic, panic, and major depressive disorders and the Five-Factor Model of personality. Journal of Nervous and Mental Disease 2001;189:154–161. [PubMed: 11277351]

Costa PT Jr, Herbst JH, McCrae RR, Siegler IC. Personality at midlife: Stability, intrinsic maturation, and response to life events. Assessment 2000;7:367–38.

Costa, PT., Jr; McCrae, RR. Revised NEO Personality Inventory (NEO-PI-R) and the NEO Five-Factor Inventory (NEO-FFI) professional manual. Odessa, FL: Psychological Assessment Resources; 1992.

Costa PT, McCrae RR, Zonderman AB. Environmental and dispositional influences on well-being: Longitudinal follow-up of an American national sample. British Journal of Psychology 1987;78:299–306. [PubMed: 3620790]

Costa, PT., Jr; Piedmont, RL. Multivariate assessment: NEO-PI-R profiles of Madeline G. In: Wiggins, JS., editor. Paradigms of personality assessment. New York: The Guilford Press; 2003. p. 262-280.

Dawson DV, Welsh-Bohmer KA, Siegler IC. Informant rated personality change in Alzheimer Disease patients: Replication, influence of premorbid profile, and covariate relationships. Research and Practice in Alzheimer's Disease 2001;5:27–32.

De Fruyt F, Bartels M, Van Leeuwen KG, De Clercq B, Decuyper M, Mervielde I. Five types of personality continuity in childhood and adolescence. Journal of Personality and Social Psychology 2006;91:538–552. [PubMed: 16938036]

De Fruyt F, Van Leeuwen KV, Bagby M, Rolland JP, Rouillon F. Assessing and interpreting personality change and continuity in patients treated for major depression. Psychological Assessment 2006;18:71–8. [PubMed: 16594814]

Eaton WW, Anthony JC, Gallo J, Cai G, Tien A, Romanoski A, Lyketsos C, Chen L. Natural history of Diagnostic Interview Schedule/DSM-IV major depression: the Baltimore ECA Followup. Archives of General Psychiatry 1997;54:993–999. [PubMed: 9366655]

Elder GH Jr. Time, human aging, and social change: Perspectives on the life course. Social Psychology Quarterly 1994;57:4–15.

Eysenck, HJ.; Eysenck, SBG. Manual of the Eysenck Personality Questionnaire. London: Hodder & Stoughton; 1975.

Folstein M, Folstein S, McHugh P. “Mini-Mental State:” A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research 1975;12:189–198. [PubMed: 1202204]

Ford DE, Vu HT, Anthony JC. Marijuana use and cessation of tobacco smoking in adults from a community sample. Drug and Alcohol Dependence 2002;67:243–248. [PubMed: 12127195]

Goldberg, LR. Language and individual differences: The search for universals in personality lexicons. In: Wheeler, L., editor. Review of personality and social psychology. Vol. 2. Beverly Hills, CA: Sage; 1981. p. 141-165.

Goldberg LR. The structure of phenotypic personality traits. American Psychologist 1993;48:26–34. [PubMed: 8427480]

Haven, S.; ten Berge, JMF. Technical Report No 29. University of Groningen; 1977. Tucker's coefficient of congruence as a measure of factorial invariance: An empirical study.

Helson R, Wink P. Personality change in women from the early 40s to the early 50s. Psychology and Aging 1992;7:46–55. [PubMed: 1558705]

Horn JL, McArdle JJ. A practical and theoretical guide to measurement invariance for aging research. Experimental Aging Research 1992;18:117–144. [PubMed: 1459160]

Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. Journal of Consulting and Clinical Psychology 1991;59:12–19. [PubMed: 2002127]

John, OP. The “Big Five” factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In: Pervin, LA., editor. Handbook of personality: Theory and research. New York: Guilford; 1990.

Johnson W, McGue M, Krueger RF. Personality stability in late adulthood: A behavioral genetic analysis. Journal of Personality 2005;73:523–551. [PubMed: 15745440]

Lanphear BP, Byrd RS, Auinger P, Schaffer SJ. Community characteristics associated with elevated blood lead levels in children. Pediatrics 1998;101:264–271. [PubMed: 9445502]

Loehlin JC, McCrae RR, Costa PT Jr, John OP. Heritabilities of common and measure-specific components of the Big Five personality factors. Journal of Research in Personality 1998;32:431–453.

Major B, O'Brien LT. The social psychology of stigma. Annual Review of Psychology 2005;56:393–421.

McCrae RR. Agreement of personality profiles across observers. Multivariate Behavioral Research 1993;28:25–4.

McCrae RR. A note on some measures of profile agreement. Journal of Personality Assessment. in press.

McCrae, RR.; Costa, PT, Jr. Personality in adulthood: A Five-Factor Theory perspective. 2nd. New York: Guilford Press; 2003.

McCrae, RR.; Costa, PT, Jr. Cross-cultural perspectives on adult personality trait development. In: Mroczek, D.; Little, T., editors. Handbook of personality development. Hillsdale, NJ: Lawrence Erlbaum Associates; 2006. p. 129-145.

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

McCrae RR, Costa PT Jr, Ostendorf F, Angleitner A, Hrebríková M, Avia MD, et al. Nature over nurture: Temperament, personality, and lifespan development. Journal of Personality and Social Psychology 2000;78:173–186. [PubMed: 10653513]

McCrae RR, Costa PT Jr, Terracciano A, Parker WD, Mills CJ, De Fruyt F, Mervielde I. Personality trait development from age 12 to age 18: Longitudinal, cross-sectional, and cross-cultural analyses. Journal Personality and Social Psychology 2002;83:1456–68.

McCrae RR, John OP. An introduction to the five-factor model and its applications. Special Issue: The five-factor model: Issues and applications. Journal of Personality 1992;60:175–215. [PubMed: 1635039]

McCrae RR, Terracciano A. 79 Members of the Personality Profiles of Cultures Project. Personality profiles of cultures: Aggregate personality traits. Journal of Personality and Social Psychology 2005;89:407–425. [PubMed: 16248722]

McCrae RR, Yik MSM, Trapnell PT, Bond MH, Paulhus DL. Interpreting personality profiles across cultures: Bilingual, acculturation and peer ratings studies of Chinese undergraduates. Journal of Personality and Social Psychology 1998;74:1041–1055. [PubMed: 9569658]

McCrae RR, Zonderman AB, Costa PT Jr, Bond MH, Paunonen SV. Evaluating replicability of factors in the Revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. Journal of Personality and Social Psychology 1996;70:552–566.

Miech RA, Eaton W, Liang K. Occupational stratification over the life course: A comparison of occupational trajectories across race and gender during the 1980s and 1990s. Work and Occupations 2003;30:440–473.

Mroczek DK, Spiro A III. Modeling intraindividual change in personality traits: Findings from the Normative Aging Study. Journal of Gerontology: Psychological Sciences 2003;58B:P153–65.

National Institute on Drug Abuse. Drug use among racial/ethnic minorities. Department of Health and Human Services; Bethesda, MD: 2003.

Paunonen SV, Zeidner M, Engvik HA, Oosterveld P, Maliphant R. The nonverbal assessment of personality in five cultures. Journal of Cross-Cultural Psychology 2000;31:220–239.

Piedmont RL. Cracking the plaster cast: Big Five personality change during intensive outpatient counseling. Journal of Research in Personality 2001;35:500–52.

Piedmont RL, McCrae RR, Riemann R, Angleitner A. On the invalidity of validity scales: Evidence from self-reports and observer ratings in volunteer samples. Journal of Personality and Social Psychology 2000;78:582–593. [PubMed: 10743882]

Roberts BW. Plaster or plasticity: Are work experiences associated with personality change in women? Journal of Personality 1997;65:205–232. [PubMed: 9226940]

Roberts BW, Caspi A, Moffitt T. The kids are alright: Growth and stability in personality development from adolescence to adulthood. Journal of Personality and Social Psychology 2001;81:670–683. [PubMed: 11642353]

Roberts BW, DelVecchio WF. The rank-order consistency of personality from childhood to old age: A quantitative review of longitudinal studies. Psychological Bulletin 2000;126:3–25. [PubMed: 10668348]

Roberts BW, Walton K, 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]

Robins RW, Caspi A, Moffitt T. It's not just who you're with, it's who you are: Personality and relationship experiences across multiple relationships. Journal of Personality 2002;70:925–964. [PubMed: 12498360]

Robins RW, Fraley RC, Roberts BW, Trzesniewski KH. A longitudinal study of personality change in young adulthood. Journal of Personality 2001;69:617–64. [PubMed: 11497032]

Samuels J, Eaton WW, Bienvenu OJ, Brown CH, Costa PT Jr, Nestadt G. Prevalence and correlates of personality disorders in the community. British Journal of Psychiatry 2002;180:536–542. [PubMed: 12042233]

Savla J, Davey A, Costa PT Jr, Whitfield KE. Replicating the NEO-PI-R factor structure in African-American older adults. Personality and Individual Differences 2007:1279–1288.

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]

Tellegen, A. Brief Manual for the Multidimensional Personality Questionnaire. University of Minnesota; Minneapolis: 1982. Unpublished manuscript

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

Terracciano A, McCrae RR, Costa PT Jr. Personality plasticity after age 30. Personality and Social Psychology Bulletin 2006;32:999–1009. [PubMed: 16861305]

Terracciano A, McCrae RR. How to measure national stereotypes? Response. Science 2006;311:777–779.

Utsey SO, Bolden MA, Lanier Y, Williams O. Examining the role of culture-specific coping as a predictor of resilient outcomes in African Americans from high-risk urban communities. Journal of Black Psychology 2007;33:75–93.

Vaidya JG, Gray EK, Haig J, Watson D. On the temporal stability of personality: Evidence for differential stability and the role of life experiences. Journal of Personality and Social Psychology 2002;83:1469–1484. [PubMed: 12500825]

Viken RJ, Rose RJ, Kaprio J, Koskenvuo M. A developmental-genetic analysis of adult personality: Extraversion and Neuroticism from 18 to 59. Journal of Personality and Social Psychology 1994;67:722–73. [PubMed: 8189349]

Widiger, TA.; Costa, PT., Jr; McCrae, RR. A proposal for Axis II: Diagnosing personality disorders using the Five-Factor Model. In: Costa, PT., Jr; Widiger, TA., editors. Personality disorders and the Five-Factor Model of personality. 2nd. Washington, DC: American Psychological Association; 2002. p. 431-456.

Wink P, Helson R. Personality change in women and their partners. Journal of Personality and Social Psychology 1993;65:597–605. [PubMed: 8410654]

**Table 1**

**Mean Scores and Internal Consistencies for NEO-PI-R Domains by Demographic Groups at Time 2**

| | | | | Mean Personality Scores | | | | | Cronbach's Alpha | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Demographic variables** | | **n** | **%** | **N** | **E** | **O** | **A** | **C** | **N** | **E** | **O** | **A** | **C** |
| Age | | | | | | | | | | | | | |
| | < 45 | 246 | 48.7 | 51.3 (10.4) | 50.0 (8.7) | 47.4 (8.6) | 48.2 (9.9) | 49.7 (9.5) | .85 | .89 | .84 | .85 | .89 |
| | ≥ 45 | 259 | 51.3 | 51.0 (9.6) | 48.0 (8.9) | 46.0 (8.5) | 51.6 (9.2) | 47.9 (9.5) | .84 | .89 | .84 | .84 | .89 |
| Gender | | | | | | | | | | | | | |
| | Female | 327 | 64.8 | 52.1 (9.9) | 49.6 (8.3) | 46.7 (8.4) | 52.2 (9.2) | 48.2 (9.6) | .84 | .88 | .84 | .84 | .88 |
| | Male | 178 | 35.2 | 49.4 (9.9) | 47.7 (9.7) | 46.6 (8.9) | 45.6 (9.2) | 49.7 (9.3) | .84 | .90 | .85 | .84 | .90 |
| Ethnicity | | | | | | | | | | | | | |
| | Black | 180 | 35.6 | 51.0 (8.9) | 48.6 (7.2) | 47.8 (7.3) | 50.0 (8.0) | 49.6 (8.7) | .79 | .88 | .79 | .79 | .88 |
| | White | 325 | 64.4 | 51.2 (10.5) | 49.2 (9.7) | 46.0 (9.2) | 49.9 (10.6) | 48.4 (9.9) | .87 | .90 | .86 | .87 | .90 |
| Education | | | | | | | | | | | | | |
| | ≤ 12 years | 284 | 56.2 | 51.5 (9.8) | 48.6 (8.6) | 44.5 (7.8) | 50.5 (9.9) | 48.0 (9.5) | .85 | .89 | .80 | .85 | .89 |
| | > 12 years | 221 | 43.8 | 48.7 (8.6) | 49.4 (9.2) | 49.4 (8.8) | 49.2 (9.4) | 49.9 (9.4) | .85 | .89 | .86 | .85 | .89 |
| Total | | 505 | 100 | 51.2 (10.0) | 49.0 (8.9) | 46.7 (8.6) | 49.9 (9.7) | 48.8 (9.5) | .85 | .89 | .84 | .85 | .89 |

*Note.* N = Neuroticism. E = Extraversion. O = Openness. A = Agreeableness. C = Conscientiousness.

**Table 2**

**Rank Order Stability Coefficients and Demographic Predictors of Rank Order Stability for Each of the NEO-PI-R Factors and Facets**

| | | Summary of Regressions for Individual Stability Scores | | | | |
|---|---|---|---|---|---|---|
| **NEO-PI-R Factor/Facet** | **Stability Coefficient** | $\beta_{Age}$ | $\beta_{Gender\ (female)}$ | $\beta_{Ethnicity\ (White)}$ | $\beta_{Education}$ | $R^2$ |
| N: Neuroticism | .68** | −.02 | −.03 | .09* | .03 | .01 |
| E: Extraversion | .70** | .10* | −.02 | .12** | .10* | .04** |
| O: Openness | .75** | .02 | −.01 | .07 | .20** | .05** |
| A: Agreeableness | .72** | .02 | −.03 | .16** | .03 | .03** |
| C: Conscientiousness | .64** | .03 | .02 | .08 | .09* | .02 |
| N1: Anxiety | .63** | −.02 | −.04 | .14** | .04 | .03** |
| N2: Angry Hostility | .57** | .02 | −.01 | .05 | .04 | .01 |
| N3: Depression | .64** | .01 | −.01 | .07 | −.02 | .01 |
| N4: Self-Consciousness | .57** | .01 | .00 | .06 | .03 | .01 |
| N5: Impulsiveness | .53** | .02 | −.02 | .09 | .07 | .01 |
| N6: Vulnerability | .64** | .03 | −.02 | .07 | .00 | .01 |
| E1: Warmth | .64** | .05 | −.06 | .08 | .10* | .03** |
| E2: Gregariousness | .60** | −.04 | −.01 | .06 | .05 | .01 |
| E3: Assertiveness | .66** | .00 | .06 | .14** | .06 | .03* |
| E4: Activity | .62** | .08 | .00 | .07 | .03 | .01 |
| E5: Excitement-Seeking | .65** | .08 | .00 | .11* | .07 | .03* |
| E6: Positive Emotions | .64** | .05 | .02 | .13** | .13** | .04** |
| O1: Fantasy | .60** | −.02 | −.02 | .08 | .07 | .01 |
| O2: Aesthetics | .67** | .02 | .00 | .07 | .18** | .04** |
| O3: Feelings | .56** | .03 | −.01 | .08 | .17** | .04** |
| O4: Actions | .59** | .03 | .02 | .13** | .14** | .04** |
| O5: Ideas | .69** | .02 | −.03 | .07 | .10* | .02 |
| O6: Values | .56** | .03 | .05 | .12** | .13** | .03** |

| NEO-PI-R Factor/Facet | Stability Coefficient | Summary of Regressions for Individual Stability Scores | | | | |
|---|---|---|---|---|---|---|
| | | $\beta_{Age}$ | $\beta_{Gender}$ (female) | $\beta_{Ethnicity}$ (White) | $\beta_{Education}$ | $R^2$ |
| A1: Trust | .63** | −.02 | .01 | .05 | .05 | .01 |
| A2: Straightforwardness | .55** | −.06 | −.09* | .10* | .07 | .03** |
| A3: Altruism | .54** | .05 | .01 | .05 | .08 | .01 |
| A4: Compliance | .58** | .07 | .00 | .10* | .05 | .02 |
| A5: Modesty | .57** | .02 | −.04 | .07 | −.01 | .01 |
| A6: Tender-Mindedness | .51** | .03 | −.09 | .09 | .05 | .02* |
| C1: Competence | .53** | .06 | −.05 | .08 | .00 | .02 |
| C2: Order | .60** | .00 | .11* | .14** | .15** | .05** |
| C3: Dutifulness | .52** | −.06 | −.09 | .03 | .03 | .02 |
| C4: Achievement Striving | .59** | .04 | −.01 | .05 | .04 | .01 |
| C5: Self-Discipline | .60** | .01 | −.01 | .08 | .05 | .01 |
| C6: Deliberation | .58** | −.01 | .04 | .05 | .05 | .01 |

*Notes*.

* $p < .05$,

** $p < .01$.

n = 505. Demographic variables were simultaneously entered as predictors of individual stability scores for each factor and facet.

## Table 3

**Paired Samples *t*-tests Examining Mean-Level Changes in NEO-PI-R Factors and Facets in the Sample as a Whole from Time 1 to Time 2**

| NEO-PI-R Factor/Facet | Time 1 | Time 2 | *d* |
|---|---|---|---|
| N: Neuroticism | 49.92 (10.25) | 51.15 (9.97) | .12** |
| E: Extraversion | 49.52 (9.20) | 48.98 (8.86) | −.06 |
| O: Openness | 47.23 (8.96) | 46.65 (8.59) | −.07* |
| A: Agreeableness | 49.82 (10.09) | 49.91 (9.71) | .01 |
| C: Conscientiousness | 48.03 (9.32) | 48.80 (9.48) | .08* |
| N1: Anxiety | 49.39 (9.56) | 51.52 (9.18) | .23** |
| N2: Angry Hostility | 50.62 (9.89) | 51.53 (9.62) | .09* |
| N3: Depression | 50.42 (10.75) | 51.36 (10.70) | .09* |
| N4: Self-Consciousness | 49.92 (9.45) | 49.66 (9.75) | −.03 |
| N5: Impulsiveness | 49.49 (9.31) | 50.05 (8.80) | .06 |
| N6: Vulnerability | 50.30 (11.18) | 51.23 (11.42) | .08* |
| E1: Warmth | 48.22 (10.02) | 47.82 (9.90) | −.04 |
| E2: Gregariousness | 50.36 (10.15) | 50.06 (9.42) | −.03 |
| E3: Assertiveness | 49.74 (9.33) | 50.11 (9.64) | .04 |
| E4: Activity | 49.07 (9.26) | 47.99 (9.70) | −.11** |
| E5: Excitement-Seeking | 48.68 (8.92) | 47.92 (9.02) | −.08* |
| E6: Positive Emotions | 49.15 (9.77) | 48.30 (10.11) | −.09* |
| O1: Fantasy | 48.68 (8.72) | 48.97 (8.32) | .03 |
| O2: Aesthetics | 49.44 (9.46) | 49.16 (9.00) | −.03 |
| O3: Feelings | 48.69 (9.30) | 48.13 (8.56) | −.06 |
| O4: Actions | 47.10 (9.62) | 45.70 (9.41) | −.15** |
| O5: Ideas | 47.35 (9.32) | 47.21 (9.36) | −.01 |
| O6: Values | 46.51 (8.96) | 45.87 (8.49) | −.07 |
| A1: Trust | 46.98 (10.60) | 46.41 (10.11) | −.06 |
| A2: Straightforwardness | 50.26 (9.63) | 50.64 (9.00) | .04 |
| A3: Altruism | 50.21 (10.20) | 50.60 (9.48) | .04 |
| A4: Compliance | 48.65 (10.17) | 48.36 (9.97) | −.03 |
| A5: Modesty | 50.12 (9.78) | 50.89 (9.37) | .08* |
| A6: Tender-Mindedness | 51.88 (10.13) | 51.56 (9.52) | −.03 |
| C1: Competence | 48.68 (10.75) | 49.30 (10.34) | .06 |
| C2: Order | 47.05 (8.66) | 47.29 (8.99) | .03 |
| C3: Dutifulness | 47.00 (9.92) | 48.01 (9.00) | .11** |
| C4: Achievement Striving | 47.91 (10.21) | 48.05 (10.57) | .01 |
| C5: Self-Discipline | 47.06 (10.47) | 47.23 (10.37) | .02 |
| C6: Deliberation | 51.36 (9.74) | 51.34 (9.92) | −.00 |

*Note*. Standard deviations are in parentheses.

$^{*}p < .05$,

$^{**}p < .01$,

$d$ = Cohen's d. n = 505

## Table 4

**Summary of ANOVAs Examining the Influence of Demographic Factors on Mean Level Changes in NEO-PI-R Factors and Facets**

| | Partial $\eta^2$ | | | |
|---|---|---|---|---|
| **NEO-PI-R Factor/Facet** | **Age × Time** | **Gender × Time** | **Ethnicity × Time** | **Education × Time** |
| N: Neuroticism | .002 | .007 | .011* | .000 |
| E: Extraversion | .002 | .001 | .001 | .002 |
| O: Openness | .002 | .000 | .008 | .008* |
| A: Agreeableness | .005 | .000 | .002 | .000 |
| C: Conscientiousness | .002 | .003 | .002 | .002 |
| N1: Anxiety | .001 | .003 | .015** | .001 |
| N2: Angry Hostility | .000 | .001 | .005 | .002 |
| N3: Depression | .000 | .008* | .014** | .007 |
| N4: Self-Consciousness | .001 | .002 | .005 | .000 |
| N5: Impulsiveness | .001 | .000 | .000 | .003 |
| N6: Vulnerability | .003 | .008* | .002 | .000 |
| E1: Warmth | .000 | .001 | .003 | .006 |
| E2: Gregariousness | .005 | .000 | .004 | .001 |
| E3: Assertiveness | .000 | .000 | .002 | .000 |
| E4: Activity | .003 | .001 | .001 | .004 |
| E5: Excitement-Seeking | .001 | .000 | .000 | .000 |
| E6: Positive Emotions | .000 | .001 | .000 | .002 |
| O1: Fantasy | .002 | .000 | .001 | .011* |
| O2: Aesthetics | .009* | .000 | .000 | .006 |
| O3: Feelings | .001 | .011* | .002 | .003 |
| O4: Actions | .006 | .002 | .002 | .004 |
| O5: Ideas | .000 | .001 | .002 | .000 |
| O6: Values | .001 | .000 | .017** | .003 |
| A1: Trust | .002 | .000 | .010* | .002 |
| A2: Straightforwardness | .000 | .000 | .000 | .000 |
| A3: Altruism | .002 | .000 | .000 | .000 |
| A4: Compliance | .002 | .000 | .004 | .000 |
| A5: Modesty | .004 | .000 | .001 | .000 |
| A6: Tender-Mindedness | .003 | .001 | .001 | .000 |
| C1: Competence | .000 | .004 | .016** | .003 |
| C2: Order | .005 | .003 | .001 | .011* |
| C3: Dutifulness | .002 | .001 | .003 | .002 |
| C4: Achievement Striving | .003 | .002 | .007 | .000 |
| C5: Self-Discipline | .005 | .007 | .011* | .003 |

| | Partial $\eta^2$ | | | |
|---|---|---|---|---|
| **NEO-PI-R Factor/Facet** | **Age × Time** | **Gender × Time** | **Ethnicity × Time** | **Education × Time** |
| C6: Deliberation | .006 | .002 | .000 | .000 |

*Notes*.

* $p < .05$,

** $p < .01$.

n = 505

## Table 5

**Percentages of Participants Showing Reliable Decrease, Increase, and Stability for Each of the NEO-PI-R Factors and Facets**

| NEO-PI-R Factor/Facet | Decrease | Stable | Increase | $\chi^2$ (2, $N$ = 505) |
|---|---|---|---|---|
| N: Neuroticism | 5.1% | 86.9% | 7.9% | 77.0** |
| E: Extraversion | 11.7% | 79.6% | 8.7% | 26.9** |
| O: Openness | 7.3% | 87.3% | 5.3% | 66.6** |
| A: Agreeableness | 8.1% | 83.8% | 8.1% | 134.3** |
| C: Conscientiousness | 8.3% | 79.8% | 11.9% | 258.4** |
| N1: Anxiety | 2.2% | 91.5% | 6.3% | 3.6** |
| N2: Angry Hostility | 4.2% | 88.5% | 7.3% | 54.9** |
| N3: Depression | 3.8% | 88.3% | 7.9% | 65.0** |
| N4: Self-Consciousness | 5.0% | 89.9% | 5.1% | 27.7** |
| N5: Impulsiveness | 3.6% | 89.9% | 6.5% | 36.6** |
| N6: Vulnerability | 7.7% | 80.6% | 11.7% | 236.5** |
| E1: Warmth | 5.1% | 89.5% | 5.3% | 32.1** |
| E2: Gregariousness | 1.3% | 83.2% | 6.5% | 163.1** |
| E3: Assertiveness | 6.2% | 86.9% | 6.9% | 73.7** |
| E4: Activity | 7.5% | 87.7% | 4.8% | 64.1** |
| E5: Excitement-Seeking | 8.1% | 85.7% | 6.1% | 95.1** |
| E6: Positive Emotions | 8.7% | 85.7% | 5.5% | 101.3** |
| O1: Fantasy | 4.2% | 91.5% | 4.4% | 13.2** |
| O2: Aesthetics | 8.1% | 85.0% | 6.9% | 108.8** |
| O3: Feelings | 4.6% | 91.9% | 3.6% | 11.3** |
| O4: Actions | 9.7% | 86.1% | 4.2% | 114.5** |
| O5: Ideas | 7.7% | 85.7% | 6.5% | 92.5** |
| O6: Values | 4.4% | 91.1% | 4.6% | 16.3** |
| A1: Trust | 7.7% | 87.3% | 5.0% | 7.4** |
| A2: Straightforwardness | 2.2% | 94.5% | 3.4% | 1.7 |
| A3: Altruism | 2.6% | 91.7% | 5.7% | 21.8** |
| A4: Compliance | 6.5% | 87.9% | 5.5% | 54.3** |
| A5: Modesty | 5.0% | 88.7% | 6.3% | 44.0** |
| A6: Tender-Mindedness | 8.5% | 83.6% | 7.9% | 139.4** |
| C1: Competence | 5.1% | 87.9% | 6.9% | 56.5** |
| C2: Order | 4.8% | 89.9% | 5.3% | 28.0** |
| C3: Dutifulness | 2.8% | 91.1% | 6.1% | 27.7** |

| NEO-PI-R Factor/Facet | Decrease | Stable | Increase | $\chi^2$ (2, $N$ = 505) |
|---|---|---|---|---|
| C4: Achievement Striving | 4.8% | 87.5% | 7.7% | 68.3** |
| C5: Self-Discipline | 6.7% | 86.5% | 6.7% | 76.2** |
| C6: Deliberation | 6.7% | 86.1% | 7.1% | 122.2** |

*Note*. The $\chi^2$-test examines deviations from the distribution expected by chance (i.e., 2.5%, 95%, 2.5%).

** $p < .01$.

$n = 505$