

Development of Personality in Early and Middle Adulthood: Set Like Plaster or Persistent Change?

Sanjay Srivastava and Oliver P. John
University of California, Berkeley

Samuel D. Gosling
University of Texas at Austin

Jeff Potter
Cambridge, Massachusetts

Different theories make different predictions about how mean levels of personality traits change in adulthood. The biological view of the Five-factor theory proposes the plaster hypothesis: All personality traits stop changing by age 30. In contrast, contextualist perspectives propose that changes should be more varied and should persist throughout adulthood. This study compared these perspectives in a large ($N = 132,515$) sample of adults aged 21–60 who completed a Big Five personality measure on the Internet. Conscientiousness and Agreeableness increased throughout early and middle adulthood at varying rates; Neuroticism declined among women but did not change among men. The variety in patterns of change suggests that the Big Five traits are complex phenomena subject to a variety of developmental influences.

How does personality change during adulthood? Psychologists since William James (1890/1950) have struggled with the question of whether various aspects of personality, including personality traits, change in meaningful ways during adulthood, and when those changes take place. Contemporary hypotheses about the development of personality traits stem from theories about what personality traits are. McCrae and Costa's (1996) five-factor theory asserts that personality traits arise exclusively from biological causes (i.e., genes) and that they reach full maturity in early adulthood; thus, this theory predicts little or no change on any personality dimension after early adulthood. By contrast, contextualist perspectives argue that traits are multiply determined, and that one important influence on traits is the individual's social environment (Haan, Millsap, & Hartka, 1986; Helson, Jones, & Kwan, 2002). Contextualist perspectives thus predict plasticity: Change is complex and ongoing, owing to the many factors that can affect personality traits.

In this study, we set out to understand how personality traits change in early and middle adulthood by examining the Big Five personality trait dimensions (Goldberg, 1992; John & Srivastava, 1999; McCrae & Costa, 1999). We used a cross-sectional design to study how mean levels of personality traits differ by age and whether those age effects are moderated by gender.¹ We were particularly interested in examining whether change on all of the Big Five dimensions stops or slows in middle adulthood, as predicted by the five-factor theory, or whether change is ongoing and differentiated, as predicted by contextualist theories.

Past Research on Mean-Level Change on the Big Five During Adulthood

A recent literature review summarized previous studies of mean-level change on the Big Five (Roberts, Robins, Caspi, & Trzesniewski, in press). In this review, Roberts et al. (in press) rationally categorized a wide variety of personality measures into the Big Five domains and summarized patterns of mean-level change that were consistent across studies. They concluded that, in general, Conscientiousness and Agreeableness tend to go up during adulthood, Neuroticism tends to go down, Openness shows mixed results across studies, and Extraversion shows no general pattern of change at the factor level. This basic pattern of findings

Sanjay Srivastava and Oliver P. John, Institute of Personality and Social Research, University of California, Berkeley; Samuel D. Gosling, Department of Psychology, University of Texas at Austin; Jeff Potter, Cambridge, Massachusetts.

Sanjay Srivastava was supported by a National Science Foundation Graduate Research Fellowship; Sanjay Srivastava and Oliver P. John were supported by National Institute of Mental Health Grant MH-43948. We thank Ravenna Helson, Robert R. McCrae, and Frank J. Sulloway for their helpful comments on drafts of this article. We also thank Frank J. Sulloway for developing the algorithm to remove repeat responders from the database.

Correspondence concerning this article should be addressed to Sanjay Srivastava, who is now at the Department of Psychology, Stanford University, Jordan Hall, Building 420, Stanford, California 94305. E-mail: sanjay@psych.stanford.edu

¹ "Change" is a broad concept that can be defined in a variety of other ways, such as rank-order change (whether people change in their ordering relative to age mates) and individual differences in change (whether different individuals change at different rates over time). These other ways of examining change address somewhat different substantive issues, and it is possible to obtain conceptually compatible but different results with the different approaches. (For a fuller discussion of different kinds of change, see Caspi & Roberts, 1999.)

has been reported in specific studies by researchers who argue that personality traits are affected by context (e.g., Helson et al., 2002; Helson & Kwan, 2000) as well as those who favor a strictly biological interpretation of traits (e.g., McCrae et al., 1999, 2000).

Although Roberts et al.'s (in press) conclusion seems to represent some common ground among researchers, there is still considerable disagreement: The biological and contextual perspectives disagree sharply over the timing of changes within the life course and over whether there are any differences between men's and women's development.

Set Like Plaster: The Five-Factor Theory

According to the five-factor theory, personality traits are "insulated from the direct effects of the environment" (McCrae & Costa, 1999, p. 144) and are exclusively biological in origin. Change is addressed by Postulate 1c of the five-factor theory: "Traits develop through childhood and reach mature form in adulthood; thereafter they are stable in cognitively intact individuals" (McCrae & Costa, 1999, p. 145). More specifically, traits are said to reach maturity by age 30 (e.g., Costa & McCrae, 1994; McCrae & Costa, 1999; McCrae et al., 2000). The predicted stability is expected to last throughout middle age, though in old age personality could change again, being disrupted by cognitive decline. A commonly used metaphor for this pattern of change, based on a passage from William James (1890/1950), is that personality becomes "set like plaster" by age 30 (see Costa & McCrae, 1994); thus, we refer to Postulate 1c, in its general form, as the *plaster hypothesis*.

In its original formulation, the plaster hypothesis stated that changes in Big Five traits after age 30 were nonexistent or trivial (Costa & McCrae, 1994; McCrae & Costa, 1990, 1996). More recently, the authors of the five-factor theory have indicated that the plaster hypothesis is "ripe for minor revision" (McCrae & Costa, 1999, p. 145), as studies have shown changes in mean levels of personality traits after age 30 (e.g., McCrae et al., 1999, 2000; see also Roberts et al., in press). They interpret such changes as stemming from intrinsic biological maturation rather than social influences, and they still regard the plaster hypothesis as basically true: "From age 18 to age 30 there are declines in Neuroticism, Extraversion, and Openness to Experience, and increases in Agreeableness and Conscientiousness; after age 30 the same trends are found, although the rate of change seems to decrease" (McCrae et al., 2000, p. 183).

Despite this conclusion, no study that we are aware of has directly tested whether mean levels of the Big Five traits do in fact change less after age 30 than before. This may be in part because past research on adult development has compared discrete age groups, rather than treating age as a continuous variable. For example, McCrae et al.'s (1999, 2000) two recent cross-sectional studies reported means for groups of 22- to 29-year-olds and means for groups of 30- to 49-year-olds, but the studies do not report the amount of change within those critical age ranges.

We thus set out to test the plaster hypothesis by directly comparing rates of change during the relevant age periods. In translating the plaster hypothesis into formal predictions about rates of change, we specified two versions of it. We call the original formulation (as described in Costa & McCrae, 1994) the *hard plaster hypothesis*: Age effects after age 30 should not be reliably different from zero, and this should hold for each of the Big Five

dimensions. We call the more recent "minor revision" (McCrae & Costa, 1999) the *soft plaster hypothesis*, because here personality is like plaster that has not fully hardened but is becoming more and more viscous: Personality traits change more slowly after age 30 than before age 30.

Contextual Perspectives on Personality and Change

In contrast to the plaster hypothesis, contextual theories predict that personality changes throughout adulthood (e.g., Haan et al., 1986; Helson, Mitchell, & Moane, 1984; Neugarten, 1972). By their very definition, contextual theories are necessarily more varied than the five-factor theory, but viewed together they predict different changes in personality during different life periods and, in some formulations, different changes for men and women (Helson, Pals, & Solomon, 1997; Wink & Helson, 1993).

Personality Changes as Person-Environment Transactions

Social roles, life events, and social environments change during the life course, and such factors have been suggested as important influences on basic personality traits (Haan et al., 1986; Hogan, 1996). A number of researchers have focused on the transactions between individuals' personalities and experiences. In the transactional view, individuals are seen as active agents who play an important role in selecting and shaping their environments, and these environments in turn affect their personalities. Often these transactions serve to amplify or strengthen earlier dispositions (Caspi & Moffitt, 1993). For example, personality traits like Openness and ambition predicted women's level of involvement in the women's movement in the 1960s and 1970s; involvement in the women's movement, in turn, led to subsequent increases in Openness and ambition (Agronick & Duncan, 1998).

Research on transactional person-environment processes generally addresses individual differences in change, but the transactional perspective can be applied to understanding mean-level change as well. Just as individual differences in personality lead individuals toward different experiences that subsequently affect their personalities, normative changes in personality help prepare people for normative adult roles, which in turn can support further personality changes. Thus, a transactional perspective on mean-level change in personality would focus on normative role transitions—that is, transitions experienced by large numbers of people.

Probably the three most important social role domains that undergo changes in early and middle adulthood are work, marriage or partnership, and parenting. These three role domains correspond to the major tasks of adulthood identified by Erikson's (1950) theory of adult development: work is involved in the adult task of consolidating an identity; marriage/partnership in the task of intimacy; and parenting children in generativity. Although individuals differ in the exact timing of when they take on work responsibilities, form committed partnerships, and nurture children, there are normative age ranges for these roles, suggesting that they may be linked to typical mean-level personality changes.

Which personality factors are related to these role domains? Conscientiousness has been linked to working, work performance, and work commitments (Barrick & Mount, 1991; Barrick, Mount, & Judge, 2001; Roberts, 1997; Vandewater & Stewart, 1998), and

to commitment to a stable partner relationship (Neyer & Asendorpf, 2001). Agreeableness should be most closely linked to parenting and similar generativity-relevant tasks, as exemplified in nurturing and prosocial behaviors (Graziano & Eisenberg, 1997; John & Srivastava, 1999). Role transitions in work, partnership, and childrearing take place throughout early and middle adulthood: Normatively, most people enter new jobs in their early 20s and begin advancing in their careers thereafter (U.S. Census Bureau, 2000a), marry in their mid to late 20s (U.S. Census Bureau, 2000b), and raise children in their 30s (U.S. Census Bureau, 2000c). If the timing of personality changes is linked to the timing of role transitions, there should be important changes in Conscientiousness and Agreeableness, and these changes should be apparent well into the 30s.

Aside from these normative social role changes, other theories suggest possible changes in personality traits after age 30. People get better at emotion regulation as they grow older and thus tend to have fewer negative emotional experiences (Gross et al., 1997); this could translate into persistently declining levels of Neuroticism with age. Socioemotional selectivity theory (Carstensen, Isaacowitz, & Charles, 1999) predicts that as adults progress into middle and later adulthood, they are less and less interested in gathering new information and in meeting new people, implying declining Openness and Extraversion, and more interested in relationships with close others, implying increasing Agreeableness.

Gender Roles and Personality Change

Do men and women differ in their development on any of the Big Five dimensions? Women and men may develop differently because of gender-based social experiences (Helson et al., 1997; Stewart & Ostrove, 1998). In particular, there may be developmental differences on Neuroticism. Adolescent girls show higher levels of Neuroticism than boys (del Barrio, Moreno-Rosset, Lopez-Martinez, & Olmedo, 1997; Gullone & Moore, 2000; Margalit & Eysenck, 1990). Yet, studies of subsequent development during middle adulthood indicate that women's self-confidence and coping skills improve with age (Helson & Moane, 1987; Helson et al., 1997), suggesting decreasing levels of Neuroticism primarily in women.

Few studies of both men and women have directly compared changes in adult men's and women's Neuroticism.² However, a large study of Finnish twins aged 18–59 followed members of multiple cohorts longitudinally and found that in both longitudinal and cross-sectional analyses, women decreased in Neuroticism with age whereas men did not change (Viken, Rose, Kaprio, & Koskenvuo, 1994). Similarly, a longitudinal study by Wink and Helson (1993) found that women became less emotionally dependent and more competent with age; in contrast, men started adulthood less dependent and more competent than women but then remained relatively stable on these traits. Thus, we expected that the gender difference in Neuroticism found in late adolescence and college-age samples would narrow with age: women should decrease in Neuroticism during adulthood, whereas men should not change much.

In summary, contextual perspectives diverge from the five-factor theory's assertion that all of the Big Five follow just one principle—no change—starting at age 30. Rather, a variety of developmental processes may affect each Big Five dimension

differently during particular life periods, possibly in different ways for men and women. Contextual perspectives, viewed together, offer a metatheoretical counterpoint to the five-factor theory: Change on the Big Five is complex and multiply determined, and remains a fact of life well beyond early adulthood. The nature of change may be different during different periods of adulthood, resulting in curvilinear age effects (Helson et al., 2002), and men and women may change in different ways, resulting in Age \times Gender interactions. Thus, we decided to examine the Big Five with regression models that would test for such differences.

Design of the Present Study

Our interest in testing hypotheses about different age effects during different developmental periods, and about different age effects for men and women, raised the issue of statistical power. Testing the hard and soft versions of the plaster hypothesis requires obtaining slope estimates for limited age ranges, and these estimates would be unreliable in small samples. Furthermore, tests of interactions and curvilinear effects have considerably less power than tests of main effects and linear trends (Chaplin, 1997; McClelland & Judd, 1993). In short, we needed a large sample to test our hypotheses. This concern led us to a medium for data collection that has been available for only a few years, but offers access to large numbers of willing participants: the Internet. The Web revolution of the mid-1990s resulted in the massive interconnection of American society to the Internet, making it possible to reach large numbers of participants. Research on Internet users indicates that, although they are not perfectly representative of the general population, they are quite diverse (Lebo, 2000; Lenhart, 2000), probably at least as much as more traditional samples of undergraduate psychology students or research volunteers recruited through newspaper advertisements or word of mouth.

Thus, we used an Internet sample to examine two issues central to adult personality development. First, in a sample sufficiently large to use new analyses well-suited to this question, we directly tested the hard and soft versions of the plaster hypothesis to see whether personality does indeed become "set like plaster" after age 30. Second, to map out the patterning of change in more detail, we used regression models with curvilinear and interactive effects to test for different changes during different life periods and for gender-specific development.

Method

Participants

Participants were part of the Gosling–Potter Internet Personality Project, a personality study of volunteers recruited and assessed over the World Wide Web. Personality and demographic data were available for 132,515 participants (54% female) between the ages of 21 and 60; the mean age of our participants was 31 years ($SD = 9$ years). All selected participants lived in the United States or Canada (the latter represented 9.2% of the sample).

² In a meta-analysis, Feingold (1994) compared the size of the gender difference between studies that used high school, college, and adult samples. Because all adult samples were grouped into a single category, however, this analysis was not sensitive to any changes in the gender gap after the age of 21.

Respondents reported their ethnicity as one of six categories: 5,710 (4.5%) respondents were Asian, 3,893 (3%) were Black, 2,414 (2%) were Latino, 2,094 (2%) were Middle Eastern, 110,004 (86%) were White, and 3,569 (3%) indicated Other; 4% of respondents declined to report their ethnicity. We re-ran all of the regression analyses using dummy coded variables for ethnicity; these controls had very little impact on the findings, and we report the analyses without these control variables.

We added a question about social class during the survey period, so this information was available for a subset of the sample. Of these 42,578 participants, 405 (1%) reported "poor," 7,614 (18%) "working class," 23,024 (54%) "middle class," 10,718 (25%) "upper-middle class," and 817 (2%) "upper class." This social-class distribution (19% below middle class, 27% above) suggests that we were including participants from a broad range of backgrounds. As with ethnicity, controlling for social class did not substantially change any of the effects reported in this article.

Procedure

The data presented here come from a noncommercial, advertisement-free Web site (www.outofservice.com) that contains personality measures as well as several games, quizzes, and questionnaires for entertainment purposes, and was publicized in a number of ways. Potential respondents could find out about the site through several channels: it could be found with major search engines under key words like *personality tests*; it was listed on portal sites, such as Yahoo!, under their directories of personality tests, and at one point was selected as a Yahoo! "Pick of the Week"; and individuals who had previously visited the Web site and signed up for its mailing list received notification when the Big Five survey was added. As is common on the Internet, news of the site apparently also spread quite widely through informal channels, such as e-mails or unsolicited links on other Web sites.

Computerized administration means that data entry and scoring are automated; thus, it is possible to recruit participants by appealing to their motivation to receive individualized personality feedback, for the purpose of self-insight or entertainment. To attract as broad and diverse a sample as possible, and to examine the effects of different recruiting approaches, we used two distinct Web pages. One was entitled "All About You—A Guide to Your Personality" and was said to measure "what many personality psychologists consider to be the fundamental dimensions of personality." The second Web page was entitled "Find Your *Star Wars* Twin," with feedback provided about "the characters from *Star Wars* with whom you are most similar (based on the Big Five personality test)." Instructions for both versions were the same and reminded participants to answer honestly to get accurate feedback. We consider later in the *Results* section how these two Web pages differed, both in who they attracted and in the substantive findings.

Measurement

The Big Five Inventory (BFI). To measure the Big Five personality dimensions, we used the BFI (John, Donahue, & Kentle, 1991). The 44 BFI items consist of short and easy-to-understand phrases to assess the prototypical traits defining each of the Big Five dimensions, making it ideal for a large survey where we could expect respondents to devote a limited amount of time. The BFI scales have shown substantial internal consistency, retest reliability, and clear factor structure, as well as considerable convergent and discriminant validity with longer Big Five measures (Benet-Martinez & John, 1998; John & Srivastava, 1999). The scales have also shown substantial agreement between self- and peer-reports (John & Paulhus, 2003; Rammstedt & John, 2003). BFI items are rated on a 5-point scale ranging from 1 = *disagree strongly* to 5 = *agree strongly* (the full BFI is reprinted in John & Srivastava, 1999).

For this report, we scored the BFI in an intuitive metric known as percentage of maximum possible (POMP) scores (P. Cohen, Cohen, Aiken,

& West, 1999). A POMP score is a linear transformation of any raw metric into a 0 to 100 scale, where 0 represents the minimum possible score and 100 represents the maximum possible score; P. Cohen et al. (1999) recommended POMP scores as a universal metric that is more intuitive than scale scores with idiosyncratic ranges. In the case of the BFI, we transformed the 1-to-5 BFI metric into POMP scores by subtracting 1 and multiplying by 25. Sample means and standard deviations, in POMP units, were as follows: Conscientiousness $M = 63.8$, $SD = 18.3$; Agreeableness $M = 66.4$, $SD = 18.0$; Neuroticism $M = 51.0$, $SD = 21.9$; Openness $M = 74.5$, $SD = 16.4$; Extraversion $M = 54.6$, $SD = 22.6$.³

Reliabilities, scale intercorrelations, and structural invariance across age groups. Reliabilities and scale intercorrelations were of special interest in this Internet sample. If there were problems with administering the BFI on the Internet, such as many random or otherwise unreliable responses, the coefficient alpha reliabilities of the five scales should be considerably lower. In contrast, attempts to self-enhance for the sake of receiving positive feedback should result in higher intercorrelations among the scales. Our results showed that neither was the case. First, the alpha reliabilities were very similar to earlier data (see John & Srivastava, 1999): $\alpha = .82$ in the Internet sample (compared with .82) for Conscientiousness; .79 (.79) for Agreeableness; .84 (.84) for Neuroticism; .80 (.81) for Openness; and .86 (.88) for Extraversion. The scale intercorrelations were also similar to previous research; the mean of the absolute discriminant correlations among the BFI scales was .16 as compared with .20 reported by John and Srivastava (1999), and the highest correlation between any two scales was only .29, as compared with .33.

Another concern was whether the BFI structure was invariant across ages; if the pattern of factor loadings was different at different ages, that would complicate the task of comparing scale scores. To test for this, we split the sample into four age groups spanning a range of 10 years each, then conducted factor analyses within each age group, extracting five factors in each analysis. The Big Five factors clearly replicated within each age group. We then computed factor congruence coefficients between age groups; the average congruence coefficient for the same conceptual factor across age groups was .99, reflecting both a high degree of structural invariance and the unusually low sampling error in such a large sample. Furthermore, we also computed the scale reliabilities separately by age and found that they did not vary with age.

Results

How did scores on the Big Five personality dimensions change with age? We report our analyses in two sections. First, we examined the two versions of the plaster hypothesis by (a) testing whether age slopes after age 30 are different from zero (hard plaster) and (b) comparing age slopes after age 30 to those before age 30 (soft plaster). Second, we tested models of the data that allow curvilinear age effects and gender differences in the magnitude of age effects, using regressions with polynomial age effects and gender interaction terms. We also estimated these models separately for the two Web page formats, to see whether effects generalized across recruitment strategies.

Is Personality Fixed After Age 30? Testing the Plaster Hypothesis

The hard plaster hypothesis asserts that there should be no age effects on any Big Five dimension after age 30; the soft plaster hypothesis asserts that age effects after age 30 should be weaker

³ Means and standard deviations broken down by age, gender, and Web page are available by request from the authors.

Table 1

Linear Slopes (and 95% Confidence Intervals) for the Age Effect on Each of the Big Five Factors, Computed Separately by Gender and Age Period

Big Five factor	Age 21–30	Age 31–60	<i>t</i> test to reject hard plaster	<i>z</i> test to confirm soft plaster	Implication of tests
Conscientiousness					
Women	.48 (±.06)	.26 (±.03)	18.4***	6.6***	Change slows but does not stop after age 30
Men	.46 (±.06)	.31 (±.04)	17.2***	4.2***	Change slows but does not stop after age 30
Agreeableness					
Women	.10 (±.06)	.28 (±.03)	21.1***	−5.4***	Change increases after age 30
Men	−.01 (±.07)	.20 (±.04)	11.1***	−5.6***	Change increases after age 30
Neuroticism					
Women	−.25 (±.07)	−.25 (±.03)	−14.5***	−0.2	Change is similar before and after age 30
Men	−.06 (±.08)	−.03 (±.04)	−1.4	−0.7	Little change before or after age 30
Openness					
Women	.04 (±.06)	−.04 (±.03)	−3.1**	−0.2	Change is of similar strength but in opposite directions before and after age 30
Men	.04 (±.05)	−.15 (±.03)	−9.4***	−3.3***	Change increases after age 30
Extraversion					
Women	.09 (±.08)	−.07 (±.04)	−4.0***	0.5	Change is of similar strength but in opposite directions before and after age 30
Men	.14 (±.08)	.05 (±.04)	2.1*	2.0*	Change slows but does not stop after age 30

Note. Sample size for age 21–30: women, $n = 41,840$; men, $n = 40,831$. Sample size for age 31–60: women, $n = 30,027$; men, $n = 19,817$. Numbers in parentheses are 95% confidence intervals. The metric for the slopes is the percentage of maximum possible units per year. Negative values of the *z* test indicate that the direction of the effect was contrary to that predicted by the soft version of the plaster hypothesis.

* $p < .05$. ** $p < .01$. *** $p < .001$.

than age effects before age 30. To test these hypotheses, we computed age slopes (i.e., regression coefficients from the Big Five dimensions regressed on age) within the two theoretically important age ranges, 21–30 and 31–60; such slopes indicate how much the predicted Big Five score increased or decreased per year. The slopes and their 95% confidence intervals are presented in Table 1.

All slopes were computed separately for men and for women. Because the plaster hypothesis makes the same predictions for men and for women, results ought to replicate across gender. Statistical tests of both versions of the plaster hypothesis are reported in Table 1. The test for the hard plaster hypothesis is the standard *t* test of whether the slope for ages 31–60 was different from zero. To test the soft plaster hypothesis, we used a *z* test (Equation 3.6.11 in J. Cohen & Cohen, 1983) to compare whether the slope for ages 21–30 was stronger than the slope for ages 31–60.⁴

Conscientiousness. As can be seen in Table 1, the raw Conscientiousness slope for men during ages 31–60 was $B = .31$, and the slope for women was $B = .26$. Both of these slopes were significantly and substantially different from zero, a clear rejection of the hard plaster hypothesis. The soft plaster hypothesis predicts that the slopes from age 31 to 60 should not be as strong as the slopes from age 21 to 30. For both men and women, the earlier slope ($B = .46$ for women and $B = .48$ for men) was stronger. The *z* tests indicated that the differences between the slopes were significant, supporting the soft plaster hypothesis for Conscientiousness. In short, this pattern indicates that people changed less in Conscientiousness after age 30 than before age 30, but they clearly did not stop changing.

Agreeableness. The results for Agreeableness disagreed sharply with both versions of the plaster hypothesis, as is evident in Table 1. Agreeableness increased significantly from age 31 to 60 for both men and women, contradicting the hard plaster hypothesis. Moreover, the age slopes for ages 31–60 were substantially greater than the age slopes for ages 21–30; that is, the data not only failed to support the soft plaster hypothesis, but they went significantly in the opposite direction.

Neuroticism. Neuroticism yielded different results for men and for women. For men, the slope for ages 31–60 was not significantly different from zero, consistent with the hard plaster hypothesis; in fact, men did not show a significant age effect in either age period. Women, however, declined consistently in Neuroticism. The slope for ages 31–60 was significantly different from zero, contradicting the hard plaster hypothesis. The age 21–30 slope for women was not significantly weaker than the age 31–60 slope, a failure to support the soft plaster hypothesis.

⁴ The hard plaster hypothesis predicts a zero effect, that is, a null hypothesis; thus, a significant result of this test would reject the hard plaster hypothesis. In contrast, the soft plaster hypothesis predicts a directional effect, that is, an alternative to a null hypothesis; thus, a significant result of this test would support the soft plaster hypothesis. The different logic of these two tests means that statistical power makes them sensitive in opposite ways. Because of the large sample size, even a very weak change after age 30 would lead to a rejection of the hard plaster hypothesis; but the large sample size also means that even a very weak soft-plaster effect would lead to a confirmation of the soft plaster hypothesis.

Openness. Openness slopes for both men and women were significantly negative after age 30, contradicting the hard plaster hypothesis. However, age effects for Openness did not differ significantly from zero for the age 21–30 decade. Because both men and women marginally increased in Openness up to age 30 and then decreased in Openness after age 30, we tested the difference of the absolute slopes to see if the magnitude was weaker after age 30. For men, the decline after age 30 was stronger than the increase up to age 30, a significant rejection of the soft plaster hypothesis. For women, the magnitude of the decline after age 30 was not significantly different from the magnitude of the increase up to age 30, a failure to support the soft plaster hypothesis.

Extraversion. Extraversion decreased significantly from age 31 to 60 for women; the increase for men from age 31 to 60 was weak and barely significant ($p = .04$) by conventional criteria. These findings contradicted the hard plaster hypothesis, though given the statistical power of the analysis, the weak result for men was not a resounding rejection. For men, the soft plaster hypothesis was confirmed for Extraversion. For women, though, the absolute strength of the increase from age 21 to 30 was not different from the absolute strength of the decrease from age 31 to 60, failing to support the soft plaster hypothesis.

Overall, then, we did not find widespread support for either version of the plaster hypothesis. Out of 10 tests of each hypothesis—five dimensions tested separately for men and women—only one fit the hard version (two if the result for men's Extraversion is counted) and only four fit the soft version.

The Pattern of Big Five Development: Modeling Age and Gender Effects

Figure 1 shows the mean scores on each of the Big Five dimensions for each age, separately for men and women. We used regression models to map out the relations of age and gender to personality. In fitting these regression models, the large sample allowed for very sensitive tests of polynomial (curvilinear) and interactive effects. However, a model selection process in which one continues to add polynomial terms as long as they meet conventional significance criteria was likely to produce unrepliable, uninterpretable, and unnecessarily complex models in such a large sample. Examining previous studies of curvilinear age effects on personality (e.g., Haan et al., 1986; Helson et al., 2002), we concluded that cubic (i.e., third-order) models with gender interactions seemed like the most complex models that could make a substantive contribution, so we set this as a practical limit on model complexity. Thus, for each Big Five factor, we considered three possible models:

$$\text{Linear: } B5 = b_0 + b_1(AGE) + b_2(GEN) + b_3(AGE*GEN).$$

$$\text{Quadratic: } B5 = b_0 + b_1(AGE) + b_2(GEN) + b_3(AGE*GEN) + b_4(AGE)^2 + b_5(AGE^2*GEN).$$

$$\text{Cubic: } B5 = b_0 + b_1(AGE) + b_2(GEN) + b_3(AGE*GEN) + b_4(AGE)^2 + b_5(AGE^2*GEN) + b_6(AGE)^3 + b_7(AGE^3*GEN).$$

(In the above equations, *B5* stands for the Big Five dimension being modeled, *AGE* represents age centered around its mean, and *GEN* is a contrast code for gender.) To select from among these

three models, we set the criterion that a more complex model would be retained only if it improved fit at $F > 25$ ($p < 10^{-5}$) over a simpler model. We used this stringent cutoff to select models that could realistically be replicated and examined further in future studies with smaller samples.

Conscientiousness. Patterns of change in work and partnership suggested that the most pronounced increases in Conscientiousness might occur during the 20s, followed by continuing growth at a slower rate. In statistical terms, this implied that the age effect on Conscientiousness would follow a decelerating function, with a steeply increasing slope in early adulthood that becomes flatter at later ages. Our decision criteria supported a quadratic model for Conscientiousness (see Table 2). A positive linear age term indicated that people were increasing in Conscientiousness at all ages, but a negative quadratic age term indicated that the rate of increase was greater at younger ages than at older ages. None of the age terms interacted with gender, indicating that men and women did not change in Conscientiousness at different rates. In Figure 1A, the quadratic functions for men and women are plotted along with the observed Conscientiousness means at each age. The close correspondence between the function and the means from the raw data suggests that the quadratic function captures the normative trend quite well.

Agreeableness. If Agreeableness is linked to nurturing and raising children, we would expect that it would increase the most during the late 20s and 30s. Because such an increase would occur in the middle of the age range being analyzed, a cubic model would be necessary to fit such an effect. The data provided clear support for a cubic model (see Table 2) under the decision criteria; the sample means and the cubic fit line are plotted in Figure 1B. The increase in Agreeableness accelerated in the late 20s, and Agreeableness continued to increase rapidly through the 30s before slowing down (but continuing to increase) in the 40s; thus, the period of most rapid increase coincided with the ages at which people are typically giving birth and nurturing their dependent children. There was a significant interaction between gender and the first-order age term, but no gender interactions with the quadratic or cubic term (see Table 2); in other words, women increased overall in Agreeableness more than men did, but the degree of curvature in the functions did not differ substantially between men and women. Though we did not predict it, this would be consistent with women, on average, being involved in nurturing roles to a greater extent than men.

Neuroticism. Neuroticism was the strongest candidate for a Gender \times Age interaction. The decision criteria indicated that Neuroticism was well described by the linear model. This model, along with the observed Neuroticism means, is graphed in Figure 1C. The significant Age \times Gender interaction (see Table 2) revealed that women declined substantially in Neuroticism throughout adulthood; men declined quite modestly. By late adulthood, the earlier gender difference had nearly disappeared.

Openness. For Openness, our model-selection criteria resulted in a linear model (see Table 2; a quadratic model would have improved the fit but only at $F[2, 132509] = 10.7$). The linear model was consistent with previous developmental findings—both men and women declined in Openness with age, but only slightly. A small, unpredicted interaction suggested that men began adulthood slightly higher in Openness but then declined at a faster rate (see Figure 1D).

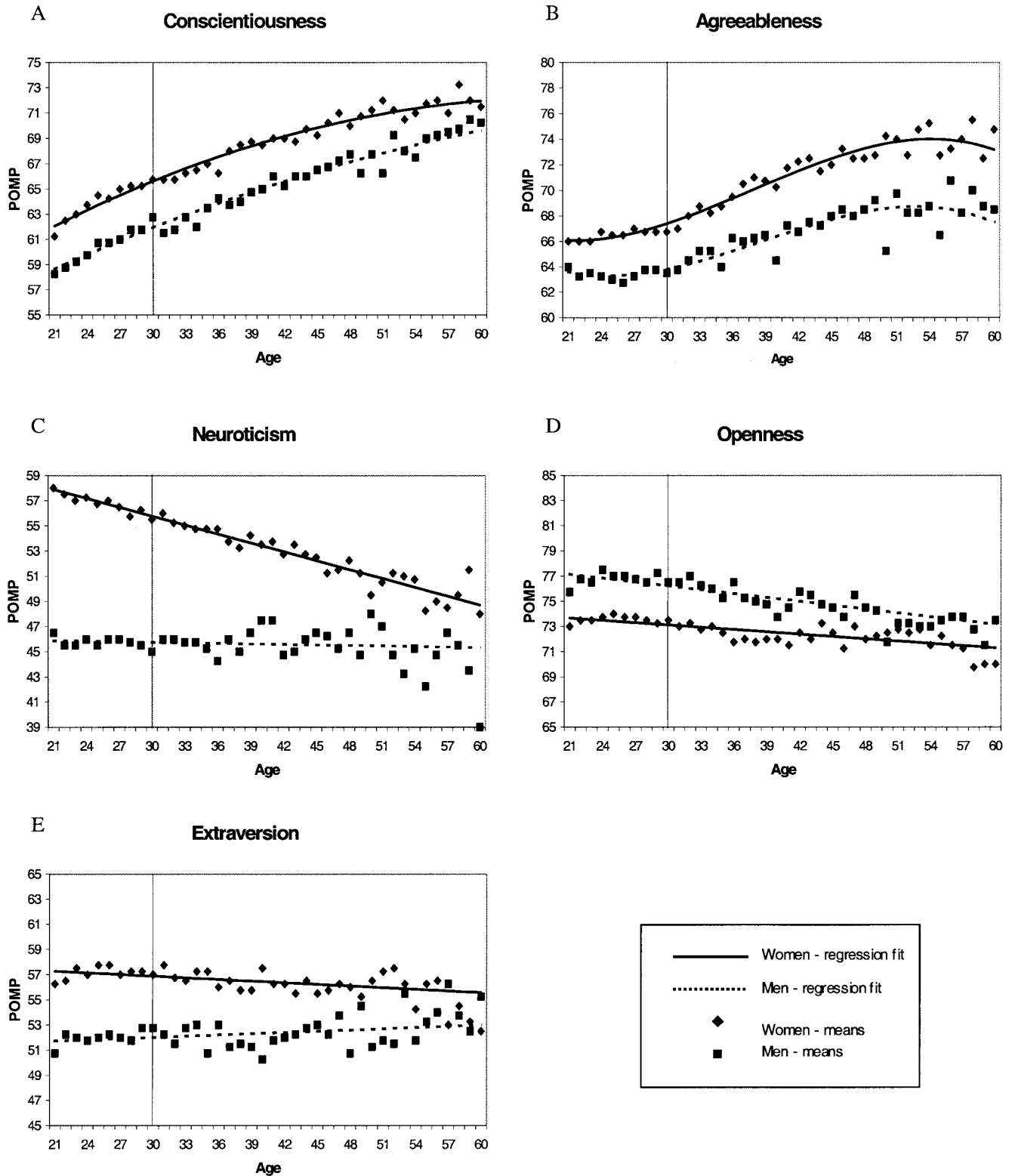


Figure 1. Mean Big Five scores broken down by age and gender, with fit curves from the regression models (see Table 2). Vertical lines at age 30 indicate when the plaster hypothesis predicts that personality stops changing. POMP = percentage of maximum possible.

Table 2
Regression Models of the Relations of Age and Gender to the Big Five

Regression term	<i>B</i>	<i>SE</i>	β	<i>p</i>	Replication β s
Conscientiousness (<i>R</i> = .18)					
Constant	63.984	.066			
Age	0.341	.008	.16	$< 10^{-22}$.19/.13
Gender	-1.847	.066	-.10	$< 10^{-22}$	-.09/-.12
Age \times Gender	-0.005	.008	-.00	$> .50$.00/.00
Age ²	-0.004	.001	-.03	$< 10^{-10}$	-.03/-.02
Age ² \times Gender	0.001	.001	.01	$> .09$.00/.01
Agreeableness (<i>R</i> = .16)					
Constant	65.735	.076			
Age	0.244	.009	.12	$< 10^{-22}$.13/.10
Gender	-1.829	.076	-.10	$< 10^{-22}$	-.10/-.10
Age \times Gender	-0.051	.009	-.03	$< 10^{-7}$	-.03/-.02
Age ²	0.011	.001	.08	$< 10^{-19}$.11/.03
Age ² \times Gender	0.001	.001	.01	$> .39$.02/.01
Age ³	-0.0005	.000	-.08	$< 10^{-17}$	-.11/-.05
Age ³ \times Gender	-0.0000	.000	-.00	$> .77$	-.01/-.01
Neuroticism (<i>R</i> = .23)					
Constant	50.690	.059			
Age	-0.125	.007	-.05	$< 10^{-22}$	-.06/-.05
Gender	-4.937	.059	-.22	$< 10^{-22}$	-.19/-.22
Age \times Gender	0.111	.007	.04	$< 10^{-22}$.05/.03
Openness (<i>R</i> = .11)					
Constant	74.628	.045			
Age	-0.082	.005	-.04	$< 10^{-22}$	-.05/-.02
Gender	1.542	.045	.09	$< 10^{-22}$.09/.04
Age \times Gender	-0.021	.005	-.01	$< 10^{-4}$	-.02/-.00
Extraversion (<i>R</i> = .11)					
Constant	54.442	.062			
Age	-0.006	.007	-.00	$> .44$.00/-.01
Gender	-2.410	.062	-.11	$< 10^{-22}$	-.10/-.11
Age \times Gender	0.038	.007	.02	$< 10^{-6}$.01/.02

Note. *N* = 132,515. Replication β s are from the All About You and Star Wars versions of the questionnaire, respectively. Age is mean-centered at 30.6. Gender is contrast-coded: female = -1, male = 1. All terms are from final models.

Extraversion. Extraversion was best fit with a linear model (see Table 2). An unpredicted Age \times Gender interaction indicated that men increased slightly in Extraversion with age whereas women decreased slightly, resulting in a diminishment of gender differences with age (see Figure 1E).

Replication of the models across Web site formats. The *p*-values associated with the tests of the polynomial models are generally quite small, some so small as to exceed the computational limits of our data analysis software (SPSS 10.0.7, which ran out of decimal places at $p < 10^{-22}$). Small *p*-values are one indicator of the reliability of an effect; cross-validation is another. The study design allowed for an internal replication analysis between the two versions of the questionnaire.

Were the respondents to the two Web pages different enough to constitute replication samples? Although we would not want to overstate the differences, the Web pages arguably appealed to somewhat different motives of potential participants. As the name implies, All About You was described as an opportunity to learn about oneself, whereas Star Wars was designed to appeal primarily as an intriguing and fun experience. The two sites did, in fact, draw somewhat different profiles of participants. All About You drew 66% women and 34% men, whereas Star Wars drew 39% women and 61% men. Respondents to All About You were on average about 2 years older than respondents to Star Wars (equivalent to an

effect size of $r = .11$). Controlling for gender, the greatest between-sites difference on the Big Five was for Openness; Star Wars respondents tended to be slightly more open to experience (partial $r = .10$).

In spite of these differences, though, the regression models clearly replicated across the two Web formats; Table 2 presents the standardized betas from these replication analyses for comparison purposes. When we analyzed the respondents for the two Web formats separately, we found that for Conscientiousness, Agreeableness, and Neuroticism, all the same terms were significant in the same direction and general magnitude, and the shape of the plotted curves was remarkably similar. For Openness, the linear term representing the predicted decline replicated across both Web formats; the unpredicted Age \times Gender term found in the full sample was not significant in the Star Wars data. For Extraversion, the main effect of gender and the Age \times Gender interaction replicated across both Web formats. These replications further underscored the reliability of the effects.

Discussion

In this article, we examined the relation between age, gender, and personality traits in adulthood. We took advantage of a large sample size and continuous age distribution to test hard and soft

versions of the plaster hypothesis, which assert that personality change stops (hard plaster) or slows (soft plaster) after age 30. On no Big Five dimension did we find support for the hard plaster hypothesis among both men and women, and only Conscientiousness demonstrated the soft plaster effect for both genders. To test for more complex change patterns, we then fitted curvilinear and interactive models to the data that yielded clear and robust effects. Conscientiousness increased throughout the age range studied, most strongly during the 20s; Agreeableness increased the most during the 30s; and Neuroticism declined with age for women but not much for men. Openness showed small declines with age, and Extraversion declined for women but not men.

Addressing Sampling and Cohort Concerns

In a study that uses a cross-sectional design to make inferences about developmental effects, differential sampling by age and cohort differences are both potential sources of confounds. Although no cross-sectional design can completely rule out such factors, they need not be fatal flaws if results are interpreted in relation to existing theory and research. In this section, we consider several factors that might address concerns about sampling and cohort confounds. We then compare age trends in the Internet sample with those found in recently published data to see whether the Internet results are comparable with those found with more traditional sampling methods and in other countries.

Sampling considerations. One concern was that Internet recruitment might produce age-confounded sampling biases.⁵ Younger Internet users might be a fairly broad group, whereas older Internet users might be a more select subset of older people in general; if so, age effects might be artifacts of who was likely to end up in our study rather than true age effects. In part, our focus on early and middle adulthood provided some assurance against such a possibility: Survey research on Internet usage and age indicates that although there is a sharp change in usage patterns around retirement age, people of preretirement age (i.e., ages 50–64) have similar access rates to younger people, and preretirement Internet users are representative of the overall Internet population in how they spend their time online (Fox et al., 2001). We were also reassured by our comparisons of the two Web page recruiting strategies. Age effects replicated across the self-learning-oriented “All About You” page and the entertainment-oriented “Find Your *Star Wars* Twin” page, suggesting that if there was a selection bias in our sampling, it would have to have been identical across recruiting strategies.

We also considered what age-confounded sampling bias would look like if it had happened. A possible scenario was that the Internet may be less familiar to older people, and thus older people may have to be higher in Openness in order to seek out and participate in an online study. In fact, we found a small drop in Openness with age, not the increase one would predict if older adults needed to be especially open in order to participate. If there was any Openness-related sampling bias, it was sufficiently weak that it did not overwhelm the age effect.⁶

Untangling developmental effects from secular trends and cohort effects. Schaie (1977) discussed three kinds of effects that developmental researchers need to consider in designing studies: (a) effects of the individual’s development, which are frequently of primary interest to developmental researchers; (b) secular trends

(i.e., changes in current social climate), which can cause variation between measurements taken at different historical times; and (c) cohort effects (i.e., generational differences), which can cause variation between people born and raised in different historical periods. Most common designs cannot fully disentangle these effects: Cross-sectional designs cannot intrinsically differentiate between developmental and cohort effects; longitudinal designs cannot separate developmental and historical effects; and time-lag designs (which compare different samples measured in different years) cannot separate historical and cohort effects. However, jointly considering different designs can provide insights. If the results of a cross-sectional study agree with results from longitudinal studies, they can be interpreted as arising from development, which is the only common effect between the two designs; and if cross-sectional results agree with results from time-lag studies, they can similarly be interpreted as arising from cohort differences. (The two interpretations are not mutually exclusive.)

In this case, the Internet findings agree with the broad trends among longitudinal studies reviewed by Roberts et al. (in press): Conscientiousness and Agreeableness went up, Neuroticism went down (their review did not differentiate men and women), and Openness and Extraversion changed little. Time-lag studies of the Big Five are far rarer than cross-sectional and longitudinal studies; two important recent meta-analyses offer a comparison source for Neuroticism (Twenge, 2000) and Extraversion (Twenge, 2001). These studies examined how, across studies, publication year correlated with sample means for anxiety and Extraversion. Both traits increased over time from the middle to late 20th century. If these results were due to cohort effects (rather than secular trends), that would show up in the cross-sectional Internet data as older people being both lower in Neuroticism and lower in Extraversion than younger people. The former was true for women but not for men, and the latter did not appear to be the case. Thus, comparisons suggest that at the level of general trends, it seems most appropriate to attribute the Internet results to developmental effects, though we cannot rule out the possibility that cohort differences may contribute additionally to explaining the age effects. It is important to note that both developmental effects and cohort effects are inconsistent with the five-factor theory; cohort effects represent an environmental influence that traits are supposed to be immune to. Thus, the age effects obtained here are inconsistent

⁵ Another potential sampling confound, not specific to an Internet design, could be selective mortality. Individuals high in Conscientiousness tend to die earlier (Friedman et al., 1993); thus, cross-sectional increases in Conscientiousness could occur because at older ages, fewer low-Conscientiousness individuals are available as participants. To test for such effects, we examined the standard deviations computed separately at each age. If age effects were due to selective mortality, the standard deviations should get smaller with increasing age, as the less conscientious participants were selected out. In fact, age was not related to the standard deviation, suggesting that the Conscientiousness effect was not driven by selective mortality.

⁶ Again, the standard deviation of Openness did not increase with age, as would be predicted by the hypothesis that the sampling procedure favored participants higher in Openness at older ages.

with the five-factor theory regardless of how such effects are substantively interpreted.

Addressing sampling and cohort effects through empirical comparisons. Another way to address concerns about this new method is through comparisons with data from studies that use more traditional approaches. If age patterns in the Internet data matched those found in cross-sectional data collected by other sampling methods, such cross-method agreement would suggest that the Internet results were not a product of sampling confounds. Furthermore, some researchers have argued that cohort effects, if they exist, should be culture-specific (Yang, McCrae, & Costa, 1998); if so, then similarities to cross-cultural findings should diminish concerns about cohort confounds.

Data in two recent articles by McCrae et al. (1999, 2000) stood out as a particularly appropriate source for comparing main effects of age. These articles, which used the Revised NEO Personality Inventory (NEO PI-R) and the NEO Five Factor Inventory (NEO-FFI) measures (Costa & McCrae, 1992), reported Big Five means for several adult age groups, allowing us to make comparisons of the direction of change and, to some extent, the approximate shape of change. Big Five self-report data were reported separately for nine different countries (i.e., Britain, Croatia, Czechoslovakia, Germany, Italy, Portugal, South Korea, Spain, and Turkey). These data had been collected under a wide variety of recruitment and sampling procedures, including a twin study, a study of parents and their teenage children, and several other convenience samples. We performed secondary analyses of the results reported from these nine samples,⁷ aggregating the *T* scores across the nine countries into averages (weighted by sample size) for the three age groups used by the original authors: 22–29, 30–49, and 50+. We were thus able to derive rough age trends (based on three data points) across the nine samples; we refer to these aggregated data as the NEO-International data. Results from data aggregated across these various cultures and recruiting procedures probably reflect a conservative number of reasonably robust effects, providing a stable baseline for making comparisons.

Although the BFI and NEO questionnaires are mostly similar, the instruments do define two of the five factors, Openness and Extraversion, somewhat differently (John & Srivastava, 1999). Thus, we also wanted to compare results with a sample that used the BFI. To do this, we derived age trends from published Big Five data collected in Germany (Lang, Luedtke, & Asendorpf, 2001). Participants from the Berlin, Germany, metropolitan area were recruited by a gold-standard procedure, stratified random sampling (stratification was on age and gender). They completed a German translation of the BFI. Only two age groups (20–40 and 45–65) in the BFI-German data overlapped with the Internet sample, so we could examine only one age interval.

Results of our secondary analyses of the NEO-International data and the BFI-German data are presented in Table 3. On Conscientiousness, both comparison samples showed increases throughout adulthood that were greater in earlier than in later adulthood, consistent with the Internet results presented in this report. Agreeableness also went up in both comparison samples; the NEO-International data showed similar-sized increases during both intervals, which is consistent with the cubic trend found in the Internet data (with only three data points, it is impossible to differentiate a cubic trend from a linear trend; this can be demonstrated by placing a straight ruler on Figure 1B such that it crosses

Table 3
Relative Magnitude of Annual Change During Two Age Intervals

Big Five dimension	Interval 1	Interval 2
Age groups that define interval		
NEO-International	22–29 vs. 30–49	30–49 vs. 50+
BFI-German	20–40 vs. 45–65	
Conscientiousness		
NEO-International	↑ ↑	↑
BFI-German	↑ ↑	
Agreeableness		
NEO-International	↑	↑
BFI-German	↑	
Neuroticism		
NEO-International	↓	↓ ↓
BFI-German	↓	
Openness		
NEO-International	↓	↓
BFI-German	0	
Extraversion		
NEO-International	↓ ↓	↓
BFI-German	0	

Note. The number of arrows indicates the relative rates of change within each row. NEO-International data come from McCrae et al. (1999, 2000); BFI-German data come from Lang et al. (2001), NEO = NEO Personality Inventory, Revised; BFI = Big Five Inventory.

the cubic curve at three points). Neuroticism declined in the comparison samples, but because the reports on the comparison samples did not provide results separately for men and women, it is not possible to know whether the trend was driven primarily by women, as it was in our sample. Openness and Extraversion declined in the NEO-International data but did not change substantially in the BFI-German; the latter is consistent with findings in the Internet data. Thus, the general direction of effects in the Internet data was similar to that of studies collected with different methods.

Nevertheless, none of the above considerations completely rule out sampling or cohort confounds, and the present study should not be taken as the last word about Big Five development in early and middle adulthood. Future longitudinal and sequential studies are needed to further address the possibility of such confounds. More importantly, we hope that such studies, which are more time-intensive and expensive than cross-sectional studies, can use the present results as a springboard to generate and test focused hypotheses about the timing and theoretical significance of changes in personality traits.

Personality Theories and Adult Development

We tested predictions about Big Five development made from two perspectives: the biologicistic Five-factor theory, and approaches that allow for contextual influences on traits. What do the results say about personality theories and adult development?

⁷ A subset of the German data reported in McCrae et al. (1999) was also included in McCrae et al. (2000). Therefore, only the 1999 data were used for the comparison baselines.

Rejection of the plaster hypothesis and its implications. Overall, we found a general lack of support for the plaster hypothesis, and considerable evidence that directly contradicted it. Mean levels of personality traits changed gradually but systematically throughout the life span, sometimes more after age 30 than before. By standardized effect size indices—multiple R s—the effects of age and gender on personality can be characterized as small to moderate (J. Cohen, 1992). So how should these findings be interpreted? The multiple R treats individual differences among participants of the same age and gender as error variance in the denominator. Thus, the R is essentially an index of predictability (Wiggins, 1973), quantifying the answer to the question, “How well can I predict an individual’s personality knowing only age and gender?” Against the backdrop of all possible predictors of personality—thousands of genes, the influence of peers and family, a lifetime of individually constructed experiences—it should not be surprising that age and gender provide at best a small to moderate amount of predictive information. In this context, the age effects reported here are indeed modest, but nevertheless constitute clear evidence of ongoing change.

The rejection of the plaster hypothesis also involves a rejection of two of its implications. The plaster hypothesis, as stated by the five-factor theory, is a general law of adult development: thus, it implies that all Big Five dimensions develop in the same ways, and that men’s and women’s development is the same. Both of these corollaries were contradicted by the evidence in this study, which showed that the course of development was different for different traits, and sometimes different for men and women. This suggests that personality traits cannot be easily explained by a small, tidy set of principles.⁸

Personality theories and traits. Perhaps ironically, personality theories that place traits at the root of everything else that is interesting about personality (McCrae & Costa, 1999) and those that consider them of little interest (Mischel & Shoda, 1995) share a common assumption: traits are separate from the rest of the personality system, static and unresponsive to social environments and life contexts. The personality system has been compared with a multistory house, with traits at the bottom and other aspects of personality—motivations, adaptations, personal narratives, and so on—built above (Little, 1996; see also McAdams, 1996). If we follow that metaphor, then the results in this article suggest what might be called a Californian view of the personality system—in California, even foundations shift a little bit year by year (and some get shaken up every once in a while).

What does this view imply for theories of personality? Strictly speaking, the evidence in this article does not by itself rule out strictly biological accounts of personality, though it does contradict the five-factor theory’s brand of biologism. Could the changes found here result solely from intrinsic, genetically determined (and sexually dimorphous) maturation? There is no denying that genes contribute to individual differences in Big Five trait levels (Loehlin, McCrae, Costa, & John, 1998). However, longitudinal behavior genetic studies, which can examine both typical levels of traits and patterns of change in traits, have suggested that change in personality traits is at least as well explained by environments as by genes, and heritabilities of personality traits diminish from early to middle adulthood (Loehlin, 1992; McGue, Bacon, & Lykken, 1993; Viken et al., 1994). Baltes (1997) referred to the “incomplete architecture of human ontogeny,” by which he meant that the

legacy of biological evolution has provided the most structure for the earlier part of the life span, when people have not yet survived long enough to pass along their genes; as individuals get older, culture becomes a more important influence on how well they adapt to their environments. Along with Baltes, we believe that adult personality is characterized more by plasticity than by increasing calcification, and the mechanisms of personality change can be understood best by considering the life contexts that accompany change.

In this study, Conscientiousness showed substantial change throughout early and middle adulthood, with the strongest effects in earlier adulthood; this is the time of life when adults are entering and advancing in the work force and forming committed partnerships, life tasks that have been linked to changes in Conscientiousness (Neyer & Asendorpf, 2001; Roberts, 1997). Agreeableness showed the largest changes somewhat later, when adults are typically caring for children. Neuroticism declined only for women, who probably had faced particularly difficult social environments earlier in their lives (Orenstein, 1994).

Several processes may link these personality and environmental changes to one another. In the case of Conscientiousness, for example, maturational changes may prompt individuals to seek out roles that fit their newly changed personalities. Conversely, new work responsibilities and relationship commitments resulting from adult social roles may require individuals to become more organized and responsible in order to meet the demands of their new situations. Or, both mechanisms may take place as part of a transactional system—personality changes may lead people to select and shape environments that further reinforce their personalities (Caspi & Roberts, 1999). Future research should further examine these processes and address more precisely what aspects of social roles and environments are transacting with personality traits. For example, are aspects of work, per se, important in the development of Conscientiousness, or is the importance of work more a matter of work’s place in broader life themes (such as being self-sufficient and taking over financial responsibilities previously assumed by parents)? Similarly, is there something specific about raising children that promotes Agreeableness (in which case the minority of adults who do not have children might prove an

⁸ The size of developmental effects has sometimes been indexed by the d statistic, expressing the difference between older and younger participants in standard deviation units. One disadvantage of this approach is that, like R , it puts all competing sources of variance (including measurement error) in the denominator; moreover, d statistics are better suited for truly categorical variables than for arbitrarily selected age groups. Nevertheless, for comparison purposes, we computed d s for each of the Big Five separately for men and women, using the difference between the highest and lowest point of each fit curve. Averaging absolute d s across all five dimensions, women changed about .33 standard deviations and men changed about .24 standard deviations. According to general scientific usage, these are medium to small effects (J. Cohen, 1992); that is, these effects do not indicate little or no change, as argued by McCrae and Costa (1996, 1999). Furthermore, averaging across the Big Five goes against an important finding for this study, namely that change differs importantly across the Big Five. In fact, we found a substantial range of effect sizes; absolute d s ranged from 0.07 (Extraversion) to 0.54 (Conscientiousness) for women, and from 0.03 (Neuroticism) to 0.59 (Conscientiousness) for men.

exception to the general finding here), or is there a broader increase in prosocial concerns in the 30s and 40s? In the case of Neuroticism, how are women able to find or create more emotionally stable social environments as they grow older?

Transactional perspectives have an important place in theories of adult development. Erikson (1950) believed that adult development was guided by a series of major challenges that grew out of biological and cultural imperatives. Identity, intimacy, and generativity—the three tasks that coincide with the age range we studied—manifest themselves in adult efforts to find work and become self-sufficient, to find a partner, and to care for others, and they are tasks that are provoked by biological imperatives but played out in the realm of adult social life and situated in a specific sociohistorical context. Neugarten (1972) theorized that biology and social environments set broad limits on adult development, within which the individual construes and creates his or her own experiences. If so, then a study of mean-level changes like this one suggests something about those broad limits, describing the typical path around which individuals vary. Roberts et al. (in press) suggested that increasing Conscientiousness and Agreeableness and decreasing Neuroticism in adulthood may indicate increasing maturity—people becoming, on average, better adapted as they get older. The findings in this article suggest that people continue to mature well into middle adulthood.

References

- Agronick, G. S., & Duncan, L. E. (1998). Personality and social change: Individual differences, life path, and importance attributed to the women's movement. *Journal of Personality and Social Psychology*, 74, 1545–1555.
- Baltes, P. B. (1997). On the incomplete architecture of human ontogeny: Selection, optimization, and compensation as foundation of developmental theory. *American Psychologist*, 52, 366–380.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1–26.
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9, 9–30.
- Benet-Martinez, V., & John, O. P. (1998). Los Cinco Grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology*, 75, 729–750.
- Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist*, 54, 165–181.
- Caspi, A., & Moffitt, T. E. (1993). When do individual differences matter? A paradoxical theory of personality coherence. *Psychological Inquiry*, 4, 247–271.
- Caspi, A., & Roberts, B. W. (1999). Personality continuity and change across the life course. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 300–326). New York: Guilford Press.
- Chaplin, W. F. (1997). Personality, interactive relations, and applied psychology. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 873–890). San Diego, CA: Academic Press.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstances for POMP. *Multivariate Behavioral Research*, 34, 315–346.
- Costa, P. T., Jr., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., & McCrae, R. R. (1994). Set like plaster: Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 21–40). Washington, DC: American Psychological Association.
- del Barrio, V., Moreno-Rosset, C., Lopez-Martinez, R., & Olmedo, M. (1997). Anxiety, depression and personality structure. *Personality and Individual Differences*, 23, 327–335.
- Erikson, E. H. (1950). *Childhood and society*. New York: Norton.
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin*, 116, 429–456.
- Fox, S., Rainie, L., Larsen, E., Horrigan, J., Lenhart, A., Spooner, T., & Carter, C. (2001). *Wired seniors: A fervent few, inspired by family ties*. Washington, DC: Pew Internet & American Life Project. Retrieved June 28, 2002, from <http://www.pewinternet.org/reports/toc.asp?Report=40>
- Friedman, H. S., Tucker, J. S., Tomlinson-Keasey, C., Schwartz, J. E., Wingard, D. L., & Criqui, M. H. (1993). Does childhood personality predict longevity? *Journal of Personality and Social Psychology*, 65, 176–185.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26–42.
- Graziano, W. G., & Eisenberg, N. (1997). Agreeableness: A dimension of personality. In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 795–824). San Diego, CA: Academic Press.
- Gross, J. J., Carstensen, L. L., Pasupathi, M., Tsai, J., Skorpen, C. G., & Hsu, A. Y. C. (1997). Emotion and aging: Experience, expression, and control. *Psychology and Aging*, 12, 590–599.
- Gullone, E., & Moore, S. (2000). Adolescent risk-taking and the five-factor model of personality. *Journal of Adolescence*, 23, 393–407.
- Haan, N., Millsap, R., & Hartka, E. (1986). As time goes by: Change and stability in personality over fifty years. *Psychology and Aging*, 1, 220–232.
- Helson, R., & Kwan, V. S. Y. (2000). Personality development in adulthood: The broad picture and processes in one longitudinal sample. In S. E. Hampson (Ed.), *Advances in personality psychology* (Vol. 1, pp. 77–106). Philadelphia: Taylor & Francis.
- Helson, R., Jones, C., & Kwan, V. S. Y. (2002). Personality change over 40 years of adulthood: Hierarchical linear modeling analyses of two longitudinal samples. *Journal of Personality and Social Psychology*, 83, 752–766.
- Helson, R., Mitchell, V., & Moane, G. (1984). Personality and patterns of adherence and nonadherence to the social clock. *Journal of Personality and Social Psychology*, 46, 1079–1096.
- Helson, R., & Moane, G. (1987). Personality change in women from college to midlife. *Journal of Personality and Social Psychology*, 53, 176–186.
- Helson, R., Pals, J. L., & Solomon, M. (1997). Is there adult development distinctive to women? In R. Hogan, J. A. Johnson, & S. R. Briggs (Eds.), *Handbook of personality psychology* (pp. 291–314). San Diego, CA: Academic Press.
- Hogan, R. (1996). A socioanalytic perspective on the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 180–207). New York: Guilford Press.
- James, W. (1950). *The principles of psychology*. New York: Dover. (Original work published 1890)
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 54*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.

- John, O. P., & Paulhus, D. L. (2003). *Self-enhancement has more than one dimension: Evidence for distinct agentic and communal biases in self-evaluation*. Unpublished manuscript.
- John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & Oliver P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 102–139). New York: Guilford Press.
- Lang, F. R., Luedtke, O., & Asendorpf, J. B. (2001). Validity and psychometric equivalence of the German version of the Big Five Inventory in young, middle-aged and old adults. *Diagnostica*, 47, 111–121.
- Lebo, H. (2000). *The UCLA internet report: Surveying the digital future*. Los Angeles: University of California Regents. Retrieved October 31, 2000, from <http://www.ccp.ucla.edu/UCLA-Internet-Report-2000.pdf>
- Lenhart, A. (2000). *Who's not online: 57% of those without Internet access say they do not plan to log on*. Washington, DC: Pew Internet & American Life Project. Retrieved October 31, 2000, from <http://www.pewinternet.org/reports/toc.asp?Report=21>
- Little, B. R. (1996). Free traits, personal projects and idio-tapes: Three tiers for personality psychology. *Psychological Inquiry*, 7, 340–344.
- Loehlin, J. C. (1992). *Genes and environment in personality development*. Newbury Park, CA: Sage.
- Loehlin, J. C., McCrae, R. R., Costa, P. T., Jr., & John, O. P. (1998). Heritabilities of common and measure-specific components of the Big Five personality factors. *Journal of Research in Personality*, 32, 431–453.
- Margalit, M., & Eysenck, S. (1990). Prediction of coherence in adolescence: Gender differences in social skills, personality, and family climate. *Journal of Research in Personality*, 24, 510–521.
- McAdams, D. P. (1996). What this framework can and cannot do. *Psychological Inquiry*, 7, 378–386.
- McClelland, G. H., & Judd, C. M. (1993). Statistical difficulties of detecting interactions and moderator effects. *Psychological Bulletin*, 114, 376–390.
- McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood*. New York: Guilford Press.
- McCrae, R. R., & Costa, P. T., Jr. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The five-factor model of personality: Theoretical perspectives* (pp. 51–87). New York: Guilford Press.
- McCrae, R. R., & Costa, P. T., Jr. (1999). A five-factor theory of personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 139–153). New York: Guilford Press.
- McCrae, R. R., Costa, P. T., de Lima, M. P., Simoes, A., Ostendorf, F., Angleitner, A., et al. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology*, 35, 466–477.
- McCrae, R. R., Costa, P. T., Ostendorf, F., Angleitner, A., Hrebickova, M., Avia, M. D., et al. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, 78, 173–186.
- McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology*, 29, 96–109.
- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review*, 102, 246–268.
- Neugarten, B. (1972). Personality and the aging process. *Gerontologist*, 12, 9–15.
- Neyer, F. J., & Asendorpf, J. B. (2001). Personality–relationship transaction in young adulthood. *Journal of Personality and Social Psychology*, 81, 1190–1204.
- Orenstein, P. (1994). *Schoolgirls: Young women, self-esteem, and the confidence gap*. New York: Doubleday.
- Rammstedt, B., & John, O. P. (2003). *Die grossen Fuenf: Measurement, language, and culture effects in Germany and the U.S.* Manuscript submitted for publication.
- Roberts, B. W. (1997). Plaster or plasticity: Are adult work experiences associated with personality change in women? *Journal of Personality*, 65, 205–232.
- Roberts, B. W., Robins, R. W., Caspi, A., & Trzesniewski, K. H. (in press). Personality trait development in adulthood. In J. L. Mortimer & M. Shanahan (Eds.), *Handbook of the life course*. New York: Plenum Press.
- Schaie, K. W. (1977). Quasi-experimental research designs in the psychology of aging. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 39–58). New York: Van Nostrand Reinhold.
- Stewart, A. J., & Ostrove, J. M. (1998). Women's personality in middle age: Gender, history, and midcourse corrections. *American Psychologist*, 53, 1185–1194.
- Twenge, J. M. (2000). The age of anxiety? The birth cohort change in anxiety and neuroticism, 1952–1993. *Journal of Personality and Social Psychology*, 79, 1007–1021.
- Twenge, J. M. (2001). Birth cohort changes in extraversion: A cross-temporal meta-analysis, 1966–1993. *Personality and Individual Differences*, 30, 735–748.
- U.S. Census Bureau. (2000a). Labor force, employment, and earnings (Section 13). *Statistical abstracts of the United States*. Washington, DC: U.S. Department of Commerce.
- U.S. Census Bureau. (2000b). Population (Section 1). *Statistical abstracts of the United States*. Washington, DC: U.S. Department of Commerce.
- U.S. Census Bureau. (2000c). Vital statistics (Section 2). *Statistical abstracts of the United States*. Washington, DC: U.S. Department of Commerce.
- Vandewater, E. A., & Stewart, A. J. (1998). Making commitments, creating lives: Linking women's roles and personality at midlife. *Psychology of Women Quarterly*, 22, 717–738.
- Viken, R. J., Rose, R. J., Kaprio, J., & Koskenvuo, M. (1994). A developmental genetic analysis of adult personality: Extraversion and neuroticism from 18 to 59 years of age. *Journal of Personality and Social Psychology*, 66, 722–730.
- Wiggins, J. S. (1973). *Personality and prediction: Principles of personality assessment*. Reading, MA: Addison-Wesley.
- Wink, P., & Helson, R. (1993). Personality change in women and their partners. *Journal of Personality and Social Psychology*, 65, 597–606.
- Yang, J., McCrae, R. R., & Costa, P. T., Jr. (1998). Adult age differences in personality traits in the United States and the People's Republic of China. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences*, 53, 375–383.

Received February 19, 2002

Revision received August 2, 2002

Accepted August 22, 2002 ■