Psychological Bulletin
2000, Vol. 126, No. 1, 3–25

0033-2909/00/$5.00 DOI: 10.1037//0033-2909.126.1.3

The Rank-Order Consistency of Personality Traits From Childhood to Old Age: A Quantitative Review of Longitudinal Studies

Brent W. Roberts and Wendy F. DelVecchio
University of Tulsa

The present study used meta-analytic techniques to test whether trait consistency maximizes and stabilizes at a specific period in the life course. From 152 longitudinal studies, 3,217 test–retest correlation coefficients were compiled. Meta-analytic estimates of mean population test–retest correlation coefficients showed that trait consistency increased from .31 in childhood to .54 during the college years, to .64 at age 30, and then reached a plateau around .74 between ages 50 and 70 when time interval was held constant at 6.7 years. Analysis of moderators of consistency showed that the longitudinal time interval had a negative relation to trait consistency and that temperament dimensions were less consistent than adult personality traits.

Do personality traits stop changing at some point during the life course? The answer to this question is critical for both basic and applied psychologists. For personality psychologists, it goes to the heart of how personality traits are conceptualized. At the core of most definitions is the assumption that traits remain consistent over time (West & Graziano, 1989). For example, Tellegen (1988) defines a trait as "a psychological (therefore) organismic structure underlying a relatively enduring behavioral disposition, i.e., a tendency to respond in certain ways under certain circumstances" (p. 622; see also Harkness & Hogan, 1995). If age is strongly related to trait consistency, then the construct validity of trait measurements may be affected by the age of the samples studied. For applied psychologists, whether traits are unchanging pertains to whether change efforts should be attempted and whether age should be factored into this decision. For example, if personality traits change before age 18 and not after, then interventions focused on changing traitlike syndromes, such as leadership style or personality disorders, may be shaped by the age of the client (see, e.g., Hellervik, Hazucha, & Schneider, 1992; Linehan & Kehrer, 1993).

Despite its obvious importance, the question of when in the life course personality traits reach their peak consistency has received little systematic empirical and quantitative attention since Bloom's (1964) review in 1964. In the present study, we focus on one aspect of personality change, rank-order consistency, by compiling longitudinal studies of personality trait consistency. We address three questions: (a) What is the relation between chronological age and trait consistency, (b) at what age does trait consistency peak, and (c) does trait consistency peak at a level high enough to warrant, as some have argued (see, e.g., McCrae & Costa, 1994), that traits stop changing at a specific age or period within the life course?

Defining Trait Consistency

To draw clear conclusions about the consistency of traits, we must distinguish among the various forms of trait consistency. Four forms appear to be the minimum number to adequately cover the domain: intraindividual differences in consistency, ipsative consistency, mean-level consistency, and rank-order consistency (Block, 1971; Block & Robins, 1993; Caspi & Roberts, 1999; Ozer, 1986). Intraindividual differences in consistency and ipsative consistency focus on whether one individual remains the same over time. The propensity for an individual to change in some magnitude or degree on a trait dimension corresponds to the study of intraindividual differences in trait consistency (Alder & Scher, 1994; Jones & Meredith, 1996; Nesselroade, 1991). The intraindividual differences approach to consistency focuses on how each individual changes with time and is most commonly operationalized by examining the correlates of difference scores (B. W. Roberts & Helson, 1997), residualized change scores (Block & Robins, 1993), or growth curve estimates of change (Tate & Hokanson, 1993). The second individual-level approach is the examination of the relative salience of attributes within an individual over time. Referred to as ipsative stability (Caspi & Roberts, 1999), this perspective on change is best exemplified by Block's (1971) research using the Q-sort technique. Block (1971) identified five male and six female patterns of ipsative change in the Berkeley Guidance and Oakland Growth longitudinal studies. For example, Block found one group of men for whom a sense of talkativeness and rebelliousness became more salient in their personalities as they moved from adolescence to young adulthood.

The two definitions of trait consistency familiar to most researchers are mean-level consistency and rank-order consistency. These definitions rely on population indexes to judge whether traits change. Mean-level consistency reflects whether groups of people increase or decrease on trait dimensions over time. If groups of people show reliable mean-level change over time, then personality is inconsistent in that it shows changes that are normative in nature. Rank-order consistency, the second population-

Brent W. Roberts and Wendy F. DelVecchio, Department of Psychology, University of Tulsa.

We thank Lewis Goldberg, Ravenna Helson, Richard Robins, Avshalom Caspi, and Daniel Bernstein for their helpful comments on earlier versions of this article. We also thank Mike Najar and Matt McClendon for their help with Table 1.

Correspondence concerning this article should be addressed to Brent W. Roberts, who is now at the Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, Illinois 61820. Electronic mail may be sent to broberts@s.psych.uiuc.edu.

level definition of trait consistency, refers to the relative placement of individuals within a group. Most commonly assessed through test–retest correlations, or stability coefficients, this perspective on consistency refers to whether groups of people retain the same rank ordering on trait dimensions over time.

The meta-analysis presented here focuses exclusively on this last definition of trait consistency. It should be noted that the existence of rank-order consistency does not rule out the possibility of other types of change such as individual-level or mean-level change (see Block, 1971; Ozer, 1986). Each methodological approach to consistency addresses a different question, and these questions are not always statistically or conceptually related. For example, rank-order consistency tracks the degree to which people change ordinal position over time (Clarke & Clarke, 1984). Whether two people change ordinal position over time is unrelated to whether they both show mean-level change (Block, 1971). Furthermore, high rank-order consistency does not rule out the possibility of individual differences in change. For example, M. L. Kohn (1980) reported that the rank-order consistency of intellectual flexibility over a 10-year period was .93 when disattenuated. Despite this remarkably high consistency, M. L. Kohn and Schooler (1978) showed that individual differences in change on intellectual flexibility in the same longitudinal sample were related to the substantive complexity of work. Given the unrelated nature of these indexes of change, we limit the inferences drawn from our study to those relevant to rank-order consistency.

Mechanisms of Personality Trait Consistency

What are the processes and mechanisms of personality trait consistency? We review five mechanisms that have been shown or hypothesized to enhance trait consistency: the environment, genes, psychological factors, person–environment transactions, and identity structure (Caspi, 1998; Caspi & Roberts, 1999).

According to numerous perspectives, a consistent environment is the most obvious yet overlooked cause of personality consistency (see, e.g., R. B. Cairns & Hood, 1983; Caspi & Roberts, 1999; Higgins & Eccles-Parsons, 1983; Moss & Susman, 1980; Sameroff, 1995). The evidence for this hypothesis tends to be indirect and focused on childhood. For example, parental child rearing practices show high levels of consistency from childhood to adolescence (McNally, Eisenberg, & Harris, 1991; G. C. Roberts, Block, & Block, 1984). Others have hypothesized that much of the consistency in adult personality traits is simply the result of living in a stable environment (see, e.g., Moss & Susman, 1980). Very few studies have tested the extent to which adult environments are consistent. In a study of the effect of work experiences on personality change, B. W. Roberts (1997) reported that status level of a person's job was more consistent than the personality trait of agency over a 16-year period (e.g., .55 vs. .42). Caspi and Herbener (1990) found that persons who married a partner similar to them were, in turn, more likely to be more consistent over time. Unfortunately, few if any of these studies directly linked environmental consistency to personality trait consistency. Glenn (1980) argued that it was not just the environment but the cumulative experience of the environment that would lead to increasing consistency in adulthood. That is, with time and age, people have fewer novel experiences. Therefore, with age and experience, people confront fewer demands to cope or adapt to environmental pressures. Tyler and Schuller (1991) reported findings that supported Glenn's hypothesis. They found that younger adults reported more change-inducing experiences than did older adults. Unfortunately, because the environment has been often overlooked in longitudinal personality research, findings supporting its effect on consistency are rare and not yet conclusive.

Genetic factors may also contribute to personality consistency. The best evidence for the role of genes in maintaining consistency was provided by McGue, Bacon, and Lykken (1993). McGue et al. administered personality tests to monozygotic and dyzygotic twins over a 10-year period. Their estimates of overall consistency were similar to other studies (ranging from .4 to .7), showing that there was a balance of consistency and change. Most interestingly, McGue et al. estimated that 80% of the personality consistency demonstrated by their sample of twins was attributable to genetic influences. Unfortunately, longitudinal twin studies of personality development are relatively rare, and no other research has replicated McGue et al.'s findings across the life course. Therefore, it is not known whether the genetic influence on consistency increases or decreases across the life course, especially in the later stages of adulthood.

A third factor that may contribute to personality trait consistency is a person's psychological make-up. That is, certain traits or cognitive structures tend to facilitate consistency across the life course. Several psychological factors associated with increased consistency cluster around the concept of adjustment and resiliency. For example, Asendorpf and Van Aken (1991) found that ego resiliency, which is, in part, related to emotional adjustment (Klohnen, 1996), predicted personality consistency over time in a longitudinal sample of children. More specifically, children who were more resilient tended to be more consistent over time. Similarly, Schuerger, Zarrella, and Hotz (1989) found that clinical samples, which one can assume are less emotionally stable, were less consistent than nonclinical samples. Finally, Clausen (1993) proposed that the trait of planful competence predicted higher levels of consistency in adulthood. People who are planfully competent tend to be more self-confident, dependable, and intellectually invested.

Several concepts theoretically related to higher levels of consistency combine both environmental and psychological factors. The first is the concept of "goodness of fit" (Thomas & Chess, 1977). According to Thomas and Chess (1977), goodness of fit results when the properties, expectations, and demands of the environment are consistent with a person's "own capacities, characteristics, and style of behaving" (p. 11). In their original conception, Thomas and Chess did not propose that goodness of fit led to consistency. Rather, they proposed that goodness of fit was associated with optimal development. Wachs (1994) proposed that goodness of fit might contribute to increased consistency in the transition from temperaments to adult personality traits. Wachs argued that children with temperaments that match their environmental characteristics might engender stabilizing reactions from their environments. In turn, these stabilizing environments would elicit a more consistent transition from temperament to personality. An elaboration on the idea of goodness of fit is the concept of developmental niche (Super & Harkness, 1994). According to Super and Harkness (1994), a niche consists of a child's physical and social settings, the customs regarding how to behave with children promoted by a community (e.g., rearing practices), and the psychology of the individuals that interact with the child. Like Wach's (1994) argument that goodness of fit may facilitate con-

sistency, the developmental niche, if consistent, may facilitate the maintenance of stable behavior patterns in the transition from childhood to adulthood.

Caspi (1998) and others (e.g., D. M. Buss, 1987; Ickes, Snyder, & Garcia, 1997) have described several types of person–environment transactions that are similar to goodness of fit and the developmental niche and are equally applicable to both children and adults (see also Caspi, Elder, & Bem, 1988; Caspi & Roberts, 1999). Like goodness of fit, person–environment transactions combine psychological and environmental factors in an explanation for why personality traits are maintained over time and context. The most widely cited types of person–environment transactions are reactive, evocative, proactive, and manipulative transactions (D. M. Buss, 1987; Caspi, 1998). Reactive transactions refer to the propensity to interpret experience in a way that is consistent with one's personality or self-concept. Evocative transactions refer to the elicitation of reactions by others that contribute to maintaining personality traits (see, e.g., Bell & Chapman, 1986). Proactive transactions refer to the propensity of a person to select roles and environments that fit best with his or her personality. Lastly, people can attempt to change their existing environments to better suit their preferences. This, reflected in attempts to change a friend or spouse's behavior, is termed manipulative. To the extent that individuals apply consciously or unconsciously reactive, evocative, proactive, and manipulative transactions, they should engender consistency in both their environments and themselves.

The final factor that may contribute to increased trait consistency is a sense of achieved or consolidated identity, which also combines features of personality and environment. Having a strong sense of identity is characterized by clarity of self, the content of which reflects psychological attributes (e.g., intellectual) and environments in the form of social roles (e.g., father; see B. W. Roberts & Donahue, 1994; Vandewater, Ostrove, & Stewart, 1997). As a cognitive schema, a strong sense of identity is both empirically and conceptually related to several earlier concepts connected to consistency. For example, identity integration has been linked to psychological adjustment and well-being (Helson, Stewart, & Ostrove, 1995). Identity consolidation, the continued investment in and evaluation of life choices made in adolescence, has been shown to predict increases in ego resiliency in young adulthood (Pals, 1999). A strong identity can also act as a potential filter of information and life experience that in turn can lead to increased consistency (i.e., similar to reactive person–environment transactions described above). In addition, an achieved or consolidated identity also lends itself to choosing life paths that are more consistent with one's personality (i.e., selective person–environment transactions). Finally, to the extent that one's identity becomes known to others in the form of a reputation (Hogan & Roberts, in press), other people may react to a person in a way that is consistent with his or her personality (i.e., evocative person–environment transactions).

In summary, we have identified environmental, genetic, psychological, and person–environment factors that all potentially contribute to higher levels of personality trait consistency with age. These factors bridge theoretical and empirical contributions from childhood through adulthood. The remaining question is the extent to which these factors may vary across the life course and thus facilitate increasing levels of trait consistency.

Age and Personality Trait Consistency

According to previous theory and research, the age at which personality traits are thought to stop changing ranges from childhood to old age. Some psychodynamic theorists claimed that personality traits were fully formed by the age of 3, mostly through childhood rearing practices (Sapir, 1934). More recently, Aldwin and Levenson (1994) claimed that personality traits were still changeable in old age. Between these two extreme positions lie perspectives outlined by both developmental and personality psychologists concerning how age relates to trait consistency, at what age trait consistency peaks, and whether trait consistency peaks at a level high enough to support the argument that personality traits stop changing.

Although childhood has seldom been considered a time when personality traits stop changing, developmental psychologists agree that personality traits proceed through several transformations in childhood that may affect consistency (see, e.g., Case, Hayward, Lewis, & Hurst, 1988; Harter, 1983). The primary transformation in childhood that should affect consistency is the transition from temperament to adult personality trait. Although the definition of temperament remains fuzzy (Goldsmith, 1996), temperaments tend to be distinguished from adult personality traits in that they are often linked directly to neurobiological functioning at birth, as well as to the early childhood environment (A. H. Buss & Plomin, 1975; Goldsmith, 1996; Rothbart & Bates, 1998; Thomas & Chess, 1977). The utility of the temperament construct, like that of personality traits, depends in part on its consistency (A. H. Buss & Plomin, 1975). The evidence for the consistency of temperament constructs ranges from relatively low to moderate levels of consistency across infancy and childhood (e.g., from 0 to .65; Kochanska, Murray, & Coy, 1997; Matheny, 1989; McDevitt, 1986).

In the broad expanse between infancy and adulthood lie gains in developmental skills that should increase temperament and trait consistency. Shiner (1998) pointed out that, in the transition to adult personality traits, temperaments most likely become more differentiated and hierarchically integrated as children age. Several of the cognitive and emotional factors associated with differentiation and integration have been identified. For example, between the ages of 3 and 5 most children develop the ability to relate their own perspective to that of someone else's (Fischer & Silvern, 1985). In addition, Sroufe (1979) proposed that children develop a self-concept during this period. Eder and Mangelsdorf (1997) reported that in middle childhood (approximately ages 6 to 12), children start to describe themselves and others with trait terms. Another developmental transition is thought to occur around age 10 or 11. Harter (1983) hypothesized that this is when children and adolescents first combine and integrate trait labels. Children at this age may also show an increased ability to move beyond simple global evaluations of self and others to more differentiated descriptions of their behavior. Similarly, Case et al. (1988) argued that adolescents acquire the use of more sophisticated defense mechanisms such as sublimation, which shows that emotions can be dealt with at a more symbolic level.

The development and increased differentiation of emotional, cognitive, and behavioral skills should lead to increasing levels of trait consistency for several reasons. For example, the development of self-conceptions provides schemas through which behaviors and actions can be evaluated. These self-conceptions can be

used by a person to interpret new events in a manner that is consistent with his or her understanding of self, especially if the actions are consistent with previously developed schemas (see Crick & Dodge, 1994; Fiske & Taylor, 1991). Also, increasing differentiation of cognitive and emotional skills permits more behaviors to be integrated into existing self-conceptions, further facilitating trait consistency.

An additional question relevant to the transition from temperaments to adult personality traits is whether temperamental differences are linked to adult differences in personality traits. Ahadi and Rothbart (1994) and others (Caspi, 1998; Digman & Shmelyov, 1996; R. P. Martin, Wisenbaker, & Huttunen, 1994; Rothbart & Bates, 1998; Shiner, 1998) have made theoretical arguments linking childhood temperaments to the adult system of personality traits known as the Big Five (e.g., extraversion, agreeableness, conscientiousness, emotional stability, and openness to experience; Goldberg, 1993). For example, both Ahadi and Rothbart and R. P. Martin et al. (1994) proposed that the temperamental trait of negative emotionality is linked to the adult trait of emotional stability. Likewise, Wachs (1994) argued that the temperamental dimension of inhibition (shyness) could be linked to both extraversion and emotional stability in adulthood (see also Digman & Shmelyov, 1996; R. P. Martin et al., 1994). Graziano, Jensen-Campbell, and Sullivan-Logan (1988) showed that activity level in childhood was linked to caregiver expectations for children's adult personality. For example, Graziano et al. found that ratings of activity level were positively associated with expectations for children to grow up extraverted. Therefore, childhood temperament may be transformed into adult personality partially through the expectancies of caregivers.

Empirical evidence linking temperaments assessed in childhood to adult differences on personality traits has been difficult to gather because of the time and effort required to track individuals from infancy to adulthood. Two recent reports provide some evidence for the temperamental basis of adult personality. The first, relying on the Dunedin Longitudinal data set (Caspi & Silva, 1995), showed, among other findings, that undercontrolled children scored lower on measures of constraint and higher on measures of negative emotion in young adulthood (age 21). Conversely, Caspi and Silva (1995) reported that inhibited children scored higher on measures of constraint and lower on measures of positive affect in adulthood. The second study, which relied on data from the Block Longitudinal Study (Block & Kremen, 1996), tested the relation between five core dimensions of temperament in childhood (activity level, task persistence, agreeableness/adaptability, negative emotionality, and social approach/withdrawal) and personality traits in adulthood (age 23) for both boys and girls (Kremen, 1999). Kremen (1999) reported that, for example, girls who were higher in approach tendencies in childhood tended to score higher on the Big Five trait of conscientiousness in adulthood, indicating that a lack of shyness and inhibition in childhood may be related to a more conventional approach to life in adulthood. Although these studies differ in the measures used in childhood and adulthood, they both show that childhood temperament is linked to adult differences in personality, albeit at a modest magnitude.

As one moves from the child development literature to the adult development literature, the research questions change from how age is related to trait consistency to when trait consistency peaks and traits stop changing. When one expects adulthood to begin is the primary marker for when trait consistency is thought to peak and traits are assumed to stop changing. Bloom (1964) hypothesized that adulthood was reached around age 20 and speculated both that personality traits reached their highest level of stability at this age and that, despite the possibility of small changes, traits were for the most part unchanging after young adulthood. Based on his review of 10 longitudinal studies of trait consistency, Bloom concluded that personality traits did not stabilize by age 20. He was unable to draw a more definitive conclusion because, at the time of his writings, only one longitudinal study of personality had followed people beyond the college years (i.e., Kelly, 1955).

A second hypothesis for when personality consistency should peak has been derived from the wealth of longitudinal research on personality development published in the past few decades (see, e.g., Conley, 1984a; Costa & McCrae, 1988; Haan, Millsap, & Hartka, 1986; Helson & Moane, 1987; B. W. Roberts & Helson, 1997). Specifically, Costa and McCrae (1988; McCrae & Costa, 1994) put forward the hypothesis that personality traits stop changing by age 30. Their inspiration, in part, comes from William James, who claimed that personality is set like plaster by age 30 (James, 1890). Based on an examination of the rank-order consistency of the Big Five trait dimensions over 3- and 6-year periods, Costa and McCrae (1988) concluded that personality traits are stable for people over age 30. Additional evidence was provided through several nonempirical overviews of longitudinal research, which contributed to the conclusion that individual differences in personality traits are fixed by age 30 (Costa & McCrae, 1997; McCrae & Costa, 1990, 1994). From this perspective, one would expect trait consistency to peak around age 30 at a level high enough to support the conclusion that traits stop changing.

The assertion that personality traits stop changing at age 30 has been questioned on both conceptual (Helson & Stewart, 1994) and empirical grounds (Field & Millsap, 1991; Helson & Wink, 1992; B. W. Roberts, 1997). Helson and Stewart (1994) criticized claims for the unchangeability of personality traits on grounds that the definition of personality was too narrow and the research strategies used to discount change were biased. Although not directly relevant to rank-order consistency, several empirical studies have reported other types of change after age 30. For example, Helson and Wink (1992) showed that women in the Mills Longitudinal Study decreased in dependence and self-criticism and increased in confidence and decisiveness between their early 40s and early 50s. In addition, Field and Millsap (1991) found increases in the Big Five dimension of agreeableness in a 14-year longitudinal study of older individuals (ages 69 to 83). In a study of individual differences in change, B. W. Roberts (1997) showed that occupational experiences, such as working in more prestigious jobs, were associated with changes in the trait dimensions of agency and norm-adherence in the transition from young adulthood to middle age. From both conceptual and empirical perspectives then, it appears that the hypothesis that personality traits stop changing at age 30 does not have uniform support.

The studies contradicting McCrae and Costa's (1994) age 30 hypothesis invite the question of whether periods in the life course beyond age 30 may be associated with higher trait consistency. To date, no theorist or researcher has pinpointed an age beyond 30 that is associated with maximal trait consistency. There are indirect indications from a variety of perspectives that the peak consistency of personality traits occurs in middle age rather than young adulthood. For example, Neugarten (1968) argued that people develop an *executive personality* in middle age. The executive personality

is characterized by increased "self-awareness, selectivity, manipulation and control of the environment, mastery, competence," and a "wide array of cognitive strategies" (p. 98). According to Neugarten, the executive personality increases one's capacity to handle complex environments and multiple pressures in both personal and interpersonal experiences. Neugarten's description of coping in middle age was supported by Vaillant's (1977) finding that defense mechanisms become increasingly sophisticated as men age. For example, Vaillant found that men decreased their use of neurotic and immature defense mechanisms, such as projection and reaction formation, and increased their use of mature defense mechanisms, such as suppression and humor, as they moved from young adulthood into middle age. Similarly, Helson and Wink (1992) found that women increased in their use of coping mechanisms such as substitution and intellectuality from age 40 to age 50. A. J. Stewart and Ostrove (1998) reported that identity certainty, which may be linked to increased personality consistency, was more prominent in middle age (ages 40 to 50) than in young adulthood. Also, Visser and Krosnick (1998) showed that attitude strength peaked in the years from 40 to 60, indicating that men and women were less likely to change their attitude in the face of persuasion during this period. Taken together, the emphasis on successful coping, integration, and certainty in these studies lends support to our alternative hypothesis that trait consistency peaks in middle age (ages 40 to 60).

Definitive evidence for when temperaments and adult personality traits shift in levels of consistency is lacking for several reasons. First, no longitudinal study has tracked individuals at numerous time points from birth to old age. Second, many of the longitudinal studies of trait consistency have only recently been published. Third, most reviews of trait consistency across the life course have been narrative overviews rather than quantitative reviews. Narrative reviews can be problematic because the conclusions drawn from them may reflect a researcher's theoretical perspective more than the data. For example, after reviewing a similar set of studies, Costa and McCrae (1997) concluded that personality traits stopped changing in young adulthood, whereas Aldwin and Levenson (1994) concluded that personality was still changeable in old age.

According to most depictions of the development of temperaments and traits, it is clear that trait consistency is assumed to increase with age. Exactly when trait consistency peaks and then stops changing is less clear. Three hypotheses have been put forward contending variably that personality consistency peaks at age 20, age 30, or in middle age, but none of these theories has received adequate attention or empirical support. We first test the hypothesis that temperaments and traits increase in consistency with age. We then test the hypotheses that trait consistency peaks at 20, 30, or in middle age, and whether the peak consistency, when reached, is close to unity.

Previous Quantitative Reviews of Personality Trait Consistency

To date, there has yet to be a comprehensive test of the relation between age and trait consistency. However, the relation between time interval and rank-order consistency of traits has been the focus of several reviews dating from World War II. We briefly review these studies because they contribute to an understanding of trait consistency and reveal some study characteristics that may affect trait consistency estimates.

Crook (1941) completed the first of these reviews. He compiled information on seven studies testing the rank-order stability of personality traits over periods as short as a few weeks (see, e.g., Neprash, 1936) and as long as 6 years (see, e.g., Crook, 1941). Crook estimated that trait consistency averaged above .80 over several weeks and dropped to around .50 after 6½ years. He also concluded that the drop was negatively accelerated, that is, the drop is fast over the initial months and then stabilizes after approximately 1 year.

Four decades later, Conley (1984a) reviewed 29 longitudinal studies of the rank-order consistency of personality. In the period since Crook's (1941) report, a number of additional reports from new longitudinal studies had been published, and many studies covered longer periods of time. The additional longitudinal studies permitted Conley to test the relation between trait consistency and time and to investigate whether specific traits, such as extraversion, neuroticism, or psychoticism, varied in their test–retest stability. Like Crook, Conley found that personality traits were more consistent over shorter time intervals. For example, when dissattenuated, measures of extraversion were quite consistent, averaging .98 over a 1-year period, approximately .70 over a 10-year period, and approximately .50 over a 40-year period. The consistency of neuroticism and psychoticism measures was, on average, lower than that of extraversion measures. Conley attributed the differences among the types of traits to differential scale reliability. Extraversion measures accumulated in Conley's study were much more internally consistent than were either neuroticism or psychoticism measures.

Schuerger et al. (1989) carried out the most comprehensive review of the rank-order consistency of traits to date. In a follow-up to an earlier review (Schuerger, Tait, & Tavernelli, 1982), Schuerger and his colleagues compiled data from 106 sources that included data on the consistency of personality traits assessed by means of eight different self-report questionnaires. Consistent with Conley's analysis, longer test–retest intervals resulted in lower rank-order consistency. In addition, nonclinical samples, or individuals who were not suffering from psychopathology, were more consistent than were clinical samples. Several scale characteristics were also predictive of trait consistency. Scale internal consistency (a combination of scale length and average interitem correlation) was predictive of higher test–retest consistency. Finally, scales from the domain of extraversion were more stable than scales assessing general adjustment (e.g., anxiety, depression). Interestingly, men and women did not differ in rank-order consistency, nor did it matter which instrument was used to assess personality.

Schuerger et al. (1989) reported that participants tended to be more consistent in their responding to personality questionnaires over later parts of the life span. Unfortunately, the effect of age on consistency was not examined explicitly, leaving the question open as to the precise point in the life course that trait consistency reaches its peak. In addition, alternative methods of assessing personality, such as observer techniques and projective methods, were not investigated by Schuerger et al. (1989) or other researchers (e.g., Conley, 1984a).

These previous studies of trait consistency should be noted for several reasons. First, the estimates of trait consistency across time have been uniformly high and relatively similar in magnitude since

Crook's (1941) early study. Second, these studies identify several factors that may enhance or undermine consistency. The most obvious factor that undermines consistency is time. Time may degrade trait consistency because of the cumulative effects of unreliability or because of the experience of true change. Furthermore, the type of trait studied appears to affect consistency estimates. Measures of the trait of extraversion appear to be more consistent than other trait domains. In addition, nonclinical samples also appear to be more consistent, implicating the role of psychological adjustment in maintaining consistent personality ordering across time. Just as important are the factors not associated with consistency. Previous studies found no gender differences in consistency and no differences across various questionnaires.

The Present Study

To address our hypotheses, we examined estimates of trait consistency drawn from longitudinal studies of temperament and adult personality traits. To better understand the relation of age and trait consistency, we categorized test–retest estimates into general age ranges associated with developmental transitions. According to Feldman (1997), the preadult years of the life course can be divided into infancy and toddlerhood (birth to age 3), the preschool period (ages 3 to 6), middle childhood (ages 6 to 12), and adolescence (ages 12 to 20). In light of the many studies of college students, we included a fifth stage from ages 18 to 21. Most reviews of the adult life course tend to use decades as important transitional periods (see, e.g., Levinson, 1978; McCrae & Costa, 1994), a practice that we followed when examining trait consistency beyond age 21. Our expectation was that trait consistency would increase with each age category until it reached a peak, most likely past age 20. In accord with prevailing theories of personality development, we expected personality to reach its peak level of consistency either in the young adult period (ages 20 to 40) or in middle age (ages 40 to 60). Once a peak level of trait consistency was identified, its proximity to unity would address our third question, which was whether trait consistency peaks at a level high enough to support the argument that personality stops changing at a given age.

In addition to testing the relation between age and trait consistency, we attempted to replicate the negative effect of time interval on trait consistency. We also tested both the assumption that increased attrition leads to higher estimates of trait consistency (Finn, 1986) and the finding that men and women do not differ in trait consistency. Previous research failed to study whether method of data collection affects estimates of trait consistency. Thus, we examined the rank-order consistency of traits across self-reports, observer ratings, and projective tests. Finally, we used emerging taxonomic systems from the temperament (R. P. Martin et al., 1994) and adult trait literature (e.g., the Big Five; Goldberg, 1993) to test whether type of trait affects rank-order consistency.

Method

Literature Searches

We used seven methods to locate studies. First, we reviewed reference lists from four quantitative reviews of rank-order consistency (Bloom, 1964; Conley, 1984a; Olweus, 1979; Scheurger et al., 1989) and two nonquantitative reviews (Aldwin & Levenson, 1994; Costa & McCrae, 1997). Second, we reviewed two databases: The first was developed by first author (Roberts), and the second was developed by Lewis Goldberg, who compiled information from 1932 to 1994 on the reliability of personality tests. Third, the PsycLIT and Dissertation Abstracts databases were searched using the following keywords: *personality consistency, personality stability, dispositional consistency, trait consistency, temperament consistency, longitudinal consistency, longitudinal stability, longitudinal temperament, and personality change. Fourth, we reviewed the Inventory of Longitudinal Studies in the Social Sciences* (Young, Savola, & Phelps, 1991). Fifth, we reviewed current issues of relevant journals (e.g., *Journal of Personality and Social Psychology, Journal of Personality, Journal of Research in Personality*). Sixth, we included databases reported in test manuals. Seventh, after developing a preliminary list of studies, we asked several knowledgeable colleagues to review the list and alert us to any studies that were overlooked.

Criteria for Study Inclusion

We included studies if they fulfilled four criteria. First, the study had to include dispositional variables (e.g., enduring, assumed consistent, cross-situational). Measures of attitudes, values, self-esteem, affect, mood, intelligence, cognitive functioning, and validity scales were not included. If these later constructs were reported in the studies compiled for the quantitative review, they were not included in the data analyses. Second, to emphasize the longitudinal consistency of traits and to diminish potential carry-over effects that could inflate estimates, we included studies with test–retest intervals greater than 1 year. Third, at a minimum, each study needed to contain information on test–retest interval, sample size, and age of the sample. Fourth, the sample studied needed to be nonclinical.

One hundred and fifty-two studies satisfied the inclusion criteria. Because many of these studies reported data from ongoing longitudinal studies, the number of samples, 124, was less than the total number of studies. The total number of participants for the 152 studies was 55,180. The total number of participants based on the 124 samples was 50,207. In all, 3,217 rank-order consistency coefficients were compiled.

Study Variables

Rank-order consistency. As described above, we examined the rank-order consistency of dispositional constructs. These included traits (e.g., extraversion), configural dispositions (e.g., narcissism), and temperaments (e.g., activity level).

Age. Age at inception of longitudinal study was coded from descriptive information given in each study. Results in a few studies were reported for a range of ages (e.g., 20–30, 30–40, etc.). For these studies, the midpoints of the reported age ranges were used as estimates of age.

For our first set of analyses focusing on the relation between age and trait consistency, we created age categories across the life course. These corresponded to stages at infancy and toddlerhood (birth to age 2.9), the preschool period (ages 3 to 5.9), middle childhood (ages 6 to 11.9), adolescence (ages 12 to 17.9), the college years (ages 18 to 21.9), and the subsequent decades through age 73. Decimal point divisions (e.g., 2.9 rather than 3) were used for the age categories through age 21.9 because some studies within these periods reported age in months rather than whole years. Studies after age 22 reported ages in years. Age 73 was used as the end age for the last age category because it represented the latest age at which a longitudinal study was initiated. Also, we combined the 60 to 70 and 70 to 80 decades because there

were too few studies that initiated longitudinal investigations after age 70.[1]

Time interval. We selected longitudinal studies that reported rank-order consistency coefficients of 1 year or longer. Interval was coded in number of years.

Attrition. Attrition was computed by subtracting the number of participants at the end of each stage of a longitudinal study from the number of participants at Time 1 and converting this figure to a percentage. Mean estimates of attrition were substituted for missing data for studies that did not report enough information to determine attrition (190 out of 3,217 coefficients, or 6% of the overall database).

Gender of sample. Gender of the sample was coded 0 = male, 1 = both male and female, and 2 = female.

Method. The method of data collection was coded into three categories: self-report, observer ratings, and projective tests. Self-report methods included paper-and-pencil approaches to assessing traits. This included standardized tests, such as the California Psychological Inventory (Gough & Bradley, 1996), scales constructed solely for research purposes, ratings of adjectives, and self-ratings of behaviors. Observer ratings included observer forms of standardized tests (see, e.g., Costa & McCrae, 1988), Q-sort rankings of traits (see, e.g., Block, 1971), ratings of childhood temperament, ratings of trait adjectives, and ratings of behaviors. Projective tests included the Rorschach inkblot test (Exner, 1980), the Thematic Apperception Test (Winter, 1988), and Loevinger's Sentence Completion test (Loevinger, 1966).

Temperament and trait categories. There appears to be an emerging consensus that temperament characteristics can be organized into five categories: approach/inhibition, adaptability, task persistence, negative emotionality, and activity level (R. P. Martin et al., 1994; Rothbart & Bates, 1998). R. P. Martin et al. (1994) also identified rhythmicity and threshold as additional categories that show up in factor analytic reviews. According to R. P. Martin et al., the existence of the rhythmicity and threshold factors may be in part the result of the wide use of the Thomas and Chess (1977) scales of temperament, which include these two dimensions. Because many longitudinal studies of temperament consistency also relied on the Thomas and Chess measures of temperament, we retained rhythmicity and threshold as temperament categories. Using definitions of the seven temperament dimensions, two independent judges categorized temperament scales into one of the seven temperament categories. Initial agreement was good (κ = .66). Discrepancies were then reviewed and discussed until consensus was reached.

The Big Five taxonomy of personality traits (Goldberg, 1993) was used to organize the personality trait test–retest coefficients into the categories of extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience. Although the Big Five subsumes many trait scales, several studies reported on dimensions outside of the Big Five. In addition to the Big Five, we coded coefficients on dimensions of femininity/masculinity and Type A. The femininity/masculinity category included measures of gender role, such as Bem's Femininity and Masculinity scales (Bem, 1974), and personality trait scales, such as Gough's measure of femininity/masculinity (Gough & Bradley, 1996). The Type A category captured the constellation of measures associated with the Type A syndrome (e.g., hostility, quickness of pace, impatience). After reviewing definitions of the Big Five and other trait coding systems (see, e.g., Hough, 1992), two judges categorized an initial list of 1,500 coefficients on the seven dimensions described above. The agreement among these judges was good (κ = .67). Discrepancies on the first 1,500 coefficients were reviewed and discussed until consensus was reached. The remaining trait coefficients were categorized by consensus.

Procedure: Aggregation of Sample Effect Sizes

Most texts on meta-analysis prescribe that the unit of analysis be the "study" and that multiple observations within the study be combined into a single average effect size (Hedges & Olkin, 1985; J. E. Hunter & Schmidt, 1990; Rosenthal, 1991). For the present series of analyses, we chose to aggregate within each sample rather than within each study because often several studies were published from one longitudinal sample.

To test the moderating effects of age and other variables on trait consistency, we developed aggregated databases from the overall database. The first database was developed to test the relation between trait consistency and age. For this age database, we aggregated sample data within the age categories described above by the age at the initiation of the longitudinal study or wave of longitudinal study in the cases where multiple assessments of the same people were performed (see, e.g., Helson & Moane, 1987). If a longitudinal study reported multiple waves of data that started within an age category (e.g., ages 21 to 22, 22 to 23, and 23 to 40), then these coefficients and the relevant study moderator variables were averaged within that age category (e.g., time and attrition). If the longitudinal study or wave of assessment spanned more than one age category, these coefficients were also averaged into the age category that was represented when the initial assessment of that study or wave of study took place. For example, in Helson and Moane (1987), the participants were contacted at ages 21, 27, and 43. Thus, there were two waves of longitudinal data. The second wave from ages 27 to 43 started in the age 22–29 category, spanned the age 30–39 category, then ended in the age 40–49 category. The data from this wave of the study were aggregated into the age 22–29 category. This technique for aggregation meant that each longitudinal sample could contribute an averaged coefficient to several separate age categories.

Similar aggregations were performed to test the relation between trait consistency and gender, method, and type of trait. Again, studies were aggregated by the potential moderator variable, such as method, and then by sample. Samples could contribute single average estimates of trait consistency and relevant moderator variables to each category within each moderator.

Analyses

To compute the estimates of trait consistency, we followed the system described by Hedges and Olkin (1985). The effect size estimates consisted of Fisher's *Z*-transformed test–retest correlation coefficients that were then weighted by the inverse of the variance when making population estimates. The estimated population correlations (ρ) for different ages were obtained through a *z*-to-*r* transformation of the effect size estimates. Confidence intervals and tests of heterogeneity were calculated using formulas from Hedges and Olkin (p. 227 and pp. 234–235, respectively). For several analyses, we also created estimates of trait consistency after controlling for relevant covariates, such as age and time interval. These estimates were made using the GLM routine of SPSS, where Fisher's *Z*-transformed test–retest correlation coefficients were weighted by the inverse of the variance, and after controlling for the relevant covariates.

Results

Study Characteristics

Table 1 shows the author, sample size, measures, types of traits represented, gender of the sample, and method used for each study in the meta-analysis. Table 1 also includes the age categories to which each study contributed estimates of trait consistency.

[1] When termination of the longitudinal assessment was used to determine the relation of age to trait consistency, the results were quite similar. We chose to use the initiation of longitudinal assessment because we planned to control for time interval in subsequent analyses. The combination of initial assessment and time interval contains all of the information in the termination of longitudinal assessment (e.g., age of initial assessment plus time interval equals age at termination) while making the effect of time interval explicit.

Table 1
Longitudinal Studies of Trait Consistency

Authors	*N*	Measures	Type of trait	Gender	Age category	Method
G. R. Adams & Fitch (1981)	148	a. Ego-Identity Incomplete Sentence Blank b. Sentence Completion Test	N, O	M&F	18–21.9	P
S. H. Adams (1994)	105	a. Cook-Medley Hostility Scale b. California Psychological Inventory Hostility Scale	A	F	18–21.9 22–29 40–49	S
Asendorpf (1990)	99	a. Parental Inhibition Scale b. Latency to Talk c. California Child Q-Sort d. Contact Initiation Coding System	App	M&F	3–5.9 6–11.9	O
Asendorpf & van Aken (1991)	238	California Child Q-Sort	N	M&F	3–5.9 6–11.9	O
Backteman & Magnusson (1981)	858	Teacher ratings of personality traits	E, A, C, N	M&F	6–11.9	O
Baltes & Nesselroade (1972)	1,249	Cattell's High School Personality Questionnaire	E, A, C, N, O	M&F	12–17.9	S
Bar-Tal & Raviv (1979)	147	a. Sociometric questionnaire of helping behavior b. Sociometric questionnaire of altruism	Ad	M&F	6–11.9	O
Bates & Pandina (1989)	1,308	Personality Research Form	E, A, C, O	M&F	12–17.9 18–21.9	S
Block (1971)	84	California Q-Sort	A	M&F	12–17.9	O
Block (1993)	102	a. California Child Q-Sort b. California Q-Sort	C, N	M&F	3–5.9 6–11.9	O
Block, Block, & Harrington (1974)	69	California Child Q-Sort	C, N	M&F	3–5.9	O
Bolton (1979)	32	16 Personality Factor Questionnaire–E. Dominance vs. Submissiveness	E, A, C, N, O	M&F	18–21.9	S
Broberg, Lamb, & Hwang (1990)	136	Interviewer ratings of temperament	Ad, App	M&F	0–2.9	O
Bromberger & Matthews (1996)	460	Beck Depression Inventory	N	F	40–49	S
Bronson (1966)	85	Interviewer ratings of personality traits	E, N	M&F	6–11.9 12–17.9	O
Bullock & Merrill (1980)	110	Peer ratings of aggression	A	M&F	6–11.9	O
D. M. Buss, Block, & Block (1980)	129	a. California Child Q-Sort b. Actometer	NA	M&F	3–5.9	O
E. Cairns, McWhirter, Duffy, & Barry (1990)	2,429	Nowicki & Strickland Locus of Control	E	M&F	12–17.9	S
Cantoni (1955)	211	Bell Adjustment Inventory	N	M&F	12–17.9 18–21.9	S
Caputo, Psathas, & Plapp (1966)	52	Edwards Personal Preference Schedule	E, A, C, N, O	F	18–21.9	S
Carmelli, Rosenman, & Chesney (1987)	370	a. Jenkins Activity Survey b. Thurstone Temperament Schedule c. Adjective Checklist Type A Scale	Type A	M	40–49	O
Carmichael & McGue (1994)	121	Eysenck Personality Inventory	E, N	M&F	12–17.9	S
Cattell & Cattell (1975)	331	Cattell's High School Personality Questionnaire	E, A, C, N, O	M&F	12–17.9	S
Conley (1984b)	441	a. Bernreuter Personality Inventory b. Bell Adjustment Inventory c. Cornell Medical Index	E, N	M&F	22–29	S
Conley (1985)	444	Bernreuter Personality Inventory	E, N	M&F	22–29	S
Cook & Wolaver (1963)[a]	322	Guilford-Zimmerman Temperament Survey	E, A, C, N, O, F/M	M&F	18–21.9	S
Costa & McCrae (1977–1978)	424	a. 16 Personality Factor Questionnaire b. Eysenck Personality Inventory	E, N	M	30–39 40–49 60–73	S
Costa & McCrae (1988)	127	a. NEO Personality Inventory b. Spouse ratings	E, A, C, N, O	M&F	40–49 60–73	S, O
Costa & McCrae (1992)	193	Guilford-Zimmerman Temperament Survey	E, A, C, N, O, F/M	F	50–59	S
Costa, McCrae, & Arenberg (1980)	410	Guilford-Zimmerman Temperament Survey	E, A, C, N, O, F/M	M	30–39 50–59 60–73	S
Crook (1941)	600	Thurstone Personality Schedule	N	F	18–21.9	O
Crook (1943)	52	Thurston Personality Schedule	N	F	18–21.9	S
M. H. Davis & Franzoi (1991)	205	a. Self-Consciousness Scale b. Interpersonal Reactivity Index	A, N, O	M&F	12–17.9	S
T. N. Davis & Satterly (1969)	149	16 Personality Factor Questionnaire	E, A, C, N, O	F	18–21.9	S
Denham, Lehman, Moser, & Reeves (1995)	38	Infant Behavior Questionnaire	NA, Pers, Ad	M&F	0–2.9	O
Digman (1989)	258	Teacher ratings of personality traits	E, A, C, N, O	M&F	6–11.9	O
Dudek & Hall (1991)	65	a. California Psychological Inventory b. Adjective Checklist	E, A, C, N, O, F/M	M	40–49	S

Table 1 (*continued*)

Authors	*N*	Measures	Type of trait	Gender	Age category	Method
Dusek & Flaherty (1981)	330	Self-ratings of "my characteristic self"	E, N, F/M	M&F	6–11.9	S
Eisenberg et al. (1987)	30	Bryant Empathy Scale	Ad	M&F	6–11.9	O
Englert (1993)	210	Parent, teacher, and self-ratings of personality	E, C, N, O	M&F	12–17.9 18–21.9 40–49	S, O
Eron, Huesmann, Lefkowitz, & Walder (1972)	427	Peer ratings of aggression	A	M&F	6–11.9	O
Exner (1980)	55	Rorschach inkblot test	N	M&F	6–11.9	P
Exner, Thomas, & Mason (1985)	57	Rorschach inkblot test	N	M&F	6–11.9 12–17.9	P
Farnsworth (1938)	55	Bernreuter Personality Inventory	E, N	M	18–21.9	S
Farrington (1978)	410	Teacher ratings of aggression	A	M	6–11.9	O
Field & Millsap (1991)	72	Interviewer ratings of personality traits	E, A, N, O	M&F	60–73	O
Finn (1986)	174	Minnesota Multiphasic Personality Inventory	E, A, C, N, O, F/M	M	18–21.9 40–49	S
Ge, Lorenz, Conger, Elder, & Simons (1994)	376	SCL-90	N	M&F	12–17.9	S
Gest (1997)	191	Observer ratings of inhibition	E	M&F	6–11.9	O
Giuganino & Hindley (1982)	97	Observer ratings of personality traits	E, N	M&F	3–5.9 6–11.9	O
Gold & Henderson (1990)	74	a. Revised Imaginal Processes Inventory b. Academic Curiosity Scale c. Revised Children's Reactive Curiosity	O	M&F	12–17.9	S
Goldsmith (1996)	37	Toddler Behavior Assessment Questionnaire	Act, NA, Pers, Ad, App	M&F	0–2.9	O
Gough (1987)	230	California Psychological Inventory	E, A, C, N, O, F/M	M&F	12–17.9	S
Gough & Bradley (1996)	328	California Psychological Inventory	E, A, C, N, O	M&F	12–17.9 18–21.9 40–49	S
Grigoriadis & Fekken (1992)		Minnesota Multiphasic Personality Inventory	E, C, N, F/M	M	30–39	S
Guerin & Gottfried (1987)	95	Minnesota Child Development Inventory	Ad, App	M&F	0–2.9	O
Guerin & Gottfried (1994)	98	a. Toddler Temperament Scale b. Behavioral Style Questionnaire	Act, NA, Pers, Ad, App, Rhy, Thresh	M&F	0–2.9 3–5.9 6–4.9	O
Haan, Millsap, & Hartka (1986)	325	California Q-Sort	E, A, C, N, O	M&F	6–11.9 12–17.9 30–39 40–49	O
Hagberg, Samuelsson, Lindberg, & Dehlin (1991)	131	Rod-and-Frame Test	O	M&F	60–73	O
Hamlin (1991)	50	Cattell's High School Personality Questionnaire	E, C, N, O	M	12–17.9	S
Harkness, Spiro, Butcher, & Ben-Porath (1995)	998	Minnesota Multiphasic Personality Inventory–2 (Psy-5)	E, A, C, N	M&F	60–73	S
Harris (1981)[b]	120	Myers-Briggs Type Indicator	E, A, C, O	M&F	22–29	S
Harsany (1993)	85	Eysenck Personality Inventory	E, N	M&F	60–73	S
Hathaway & Monachesi (1963)	3,976	Minnesota Multiphasic Personality Inventory	E, C, N, F/M	M&F	12–17.9	S
Heinicke, Diskin, Ramsey-Klee, & Oates (1986)	44	Observer ratings of attentiveness	Pers	M&F	0–2.9	O
Helson & Moane (1987)	81	California Psychological Inventory	E, A, C, N, O, F/M	F	18–21.9 22–29	S
Helson, Roberts, & Agronick (1995)	104	California Psychological Inventory	O	F	18–21.9	S
Helson & Wink (1992)	96	Adjective Checklist	E, A, C, N, O, F/M	F	40–49	S
Holmlund (1991)	349	Cesarec-Marke Personality Schedule	E, A, C, O	F	12–17.9	S
Holmlund (1992)	349	Cesarec-Marke Personality Schedule	E, A, C, N, O	F	12–17.9	S
Huesmann, Eron, Lefkowitz, & Walder (1984)	427	Peer ratings of aggression	A	M&F	6–11.9	S
S. M. Hunter, Johnson, Vizelberg, Webber, & Berenson (1991)	1,744	Hunter-Wolf Type A Behavior Pattern	E, A, C, N, Type A	M	6–11.9 18–21.9 22–29	S
Jessor (1983)	595	Personality System	A, C, N, O, F/M	M&F	12–17.9 22–29	S
John, Cheek, & Klohnen (1996)	82	California Q-Sort	E	M&F	18–21.9	O
Kagan (1960)	63	Rorschach inkblot test	N	M&F	6–11.9	P
Kagan & Moss (1962)	89	Observer ratings of behaviors	NA, Pers, Ad, App	M&F	0–2.9 3–5.9 6–11.9	O

(*table continues*)

Table 1 (*continued*)

Authors	*N*	Measures	Type of trait	Gender	Age category	Method
Kelly (1955)	368	Bernreuter Personality Inventory	E	M&F	22–29	S
Keltikangas-Jarvinen (1989)	1,314	AFMS questionnaire	Type A	M&F	12–17.9 18–21.9	S
Kochanska, Murray, & Coy (1997)	83	a. Observer ratings of inhibitory control b. Test data c. Children's Behavior Questionnaire	Pers	M&F	0–2.9 3–5.9	O
Kochanska, Murray, Jacques, Koenig, & Vandergeest (1996)	99	a. Observer ratings of inhibitory control b. Test data	Pers	M&F	0–2.9	O
M. Kohn & Rosman (1972)	486	a. Kohn Social Competence Scale b. Kohn Symptom Checklist	E, C	M&F	3–5.9	O
M. Kohn & Rosman (1973)	271	a. Kohn Social Competence Scale b. Kohn Symptom Checklist		M	3–5.9	O
Korn (1984)	127	New York Longitudinal Study Temperament Scales	Ad	M&F	0–2.9 3–5.9	O
Leon, Gillum, Gillum, & Gouze (1979)	71	Minnesota Multiphasic Personality Inventory	E, C, N, F/M	M	40–49 50–59 60–73	S
Lerner, Hertzog, Hooker, Hassibi, & Thomas (1988)	75	Psychiatrist's ratings of negative emotionality and aggression	NA, Ad	M&F	3–5.9	O
Loehlin, Horn, & Willerman (1990)	312	Parent ratings of personality	E, C, N	M&F	6–11.9	O
Loevinger et al. (1985)	648	Washington University Sentence Completion	O	M	18–21.9	S
Lovibond (1998)	882	Depression Anxiety Stress Scales	N	M&F	18–21.9	S
Magnusson & Backteman (1978)	788	a. Consequence & Divergent Figures b. Brick Uses c. Plot Titles d. Purdue Creativity Test	O	M&F	12–17.9	S
J. Martin & Redmore (1978)	32	Washington University Sentence Completion	O	M&F	6–11.9	P
Masten, Morison, & Pellegrini (1985)	163	Revised Class Play	E, A	M&F	6–11.9	O
Matheny (1983)	291	Infant Behavior Record	Act, Pers, App	M&F	0–2.9	O
Matheny (1989)	130	a. Lab observations b. Infant Behavior Record c. Toddler Temperament Scale	NA, App	M&F	0–2.9	O
McDevitt & Carey (1981)	115	Toddler Temperament Scale	All scales	M&F	0–2.9	O
McGue, Bacon, & Lykken (1993)	254	Multidimensional Personality Questionnaire	E, A, C, N, O	M&F	18–21.9	S
McNeil & Persson-Blennow (1988)	160	New York Longitudinal Study rating scales	Act, NA, Pers, Ad, App, Rhy, Thresh	M&F	0–2.9	O
Melamed, Silverman, & Lewis (1974)	62	16 Personality Factor Questionnaire	E, A, C, N, O	F	30–39	S
Meyer, Heath, Eaves, Mosteller, & Schieken (1988)	100	Cattell's Children's Personality Questionnaire	E, A, C, N, O	M&F	6–11.9	S
Mortimer, Finch, & Kumka (1982)	368	Semantic differential scale	E, C, N, O	M	18–21.9	S
Muntaner, Garcia-Sevilla, Fernandez, & Torrubia (1988)	29	a. Claridge & Brooks' Structure of Temperament Questionnaire b. Eysenck Personality Inventory	E, C, N	M&F	18–21.9	O
Mussen, Eichorn, Honzik, Bieber, & Meredith (1980)	81	Observer ratings of behavior	E, A, N, O	F	30–39	O
Myers (1973)[c]	203	Myers-Briggs Type Indicator	E, A, C, O	M&F	12–17.9 18–21.9	S
Nichols (1967)	636	16 Personality Factor Questionnaire	E, C, N, N, O, F/M	M&F	18–21.9	S
Nolen-Hoeksema, Girgus, & Seligman (1986)	139	a. Children's Depression Inventory b. Children's Attributional Style Questionnaire	N	M&F	6–11.9	S
O'Donnell, Leicht, Phillips, Mamett, & Horn (1988)	164	Behavior Problem Checklist	E, C, N	M&F	6–11.9	O
Ogawa, Sroufe, Weinfield, Carlson, & Egeland (1997)	163	Dissociation composite scores	N	M&F	3–5.9 6–11.9 12–17.9	O
Olson (1989)	50	a. Delay of gratification b. Kansas Reflection Impulsivity Scale	Pers	M&F	3–5.9	O
Olweus (1977)	85	Peer ratings of aggression	A	M	12–17.9	O
Ormel & Schaufeli (1991)	615	Andriessen & Van Cadsen Locus of Control Scale	E	M&F	22–29 40–49	S
Pederson (1991)	553	a. Zuckerman Sensation Seeking b. General Health Questionnaire	E, C, N, O	M&F	12–17.9	S
Pedlow, Sanson, Prior, & Oberklaid (1993)	450	a. Revised Infant Temperament Questionnaire b. Toddler Temperament Scale c. Childhood Temperament Questionnaire	NA, Pers, Ad, App, Rhy	M&F	0–2.9 3–5.9	O

Table 1 (*continued*)

Authors	*N*	Measures	Type of trait	Gender	Age category	Method
Piccione, Hilgard, & Zimbardo (1989)	50	Stanford Hypnotic Susceptibility Scale	O	M&F	18–21.9	O
Plant (1965a)	2,151	Modified California Ethnocentrism Scale	O	M&F	18–21.9	S
Plant (1965b)	2,890	Dogmatism Scale	O	M&F	18–21.9	S
Plant & Telford (1966)	1,713	California Psychological Inventory	E, C, O	M&F	18–21.9	S
Popham & Holden (1991)	55	Minnesota Multiphasic Personality Inventory	E, A, N, O, F/M	M&F	18–21.9	S
Redmore (1983)	97	Washington University Sentence Completion	O	M&F	18–21.9	P
Redmore & Loevinger (1979)	442	Washington University Sentence Completion	O	M&F	6–11.9	P
Reznick, Gibbons, Johnson, & McDonough (1989)	76	Composite of indicators of inhibition	App	M&F	0–2.9	O
B. W. Roberts & Chapman (in press)	77	California Psychological Inventory	N	F	18–21.9 22–29 40–49	S
Rubin, Hymel, & Mills (1989)	69	Peer and teacher ratings of sociability and social withdrawal	E	M&F	6–11.9	O
Ruff, Lawson, Parrinello, & Weissberg (1990)	154	a. Observer ratings and test data on sustained attention b. Conner's Parent Questionnaire	Pers	M&F	0–2.9	O
Sanson, Pedlow, Cann, Prior, & Oberklaid (1996)	501	a. Revised Infant Temperament Questionnaire b. Toddler Temperament Scale c. Childhood Temperament Scale	App	M&F	0–2.9 3–59	O
Saudino & Eaton (1995)	106	a. Actometer b. Infant Behavior Questionnaire c. Toddler Behavioral Assessment Questionnaire	Act	M&F	0–2.9	O
Schofield (1953)	83	Minnesota Multiphasic Personality Inventory	C, N, F/M	M	18–21.9	S
Skolnick (1966)	91	Thematic Apperception Test	E, A, C	M&F	12–17.9	P
Soldz & Vaillant (1999)	163	Ratings of Big Five and NEO Personality Inventory	E, A, C, N, O	M	22–29	S&O
Stacy, Newcomb, & Bentler (1991)	584	Zuckerman Sensation Seeking	E	M&F	18–21.9	O
Stein, Newcomb, & Bentler (1986)	654	Bentler Psychological Inventory	E, A, C, N, O	M&F	12–17.9 18–21.9	S
Steinberg (1986)	73	a. Teacher interviews b. Matthew's Youth Test of Health	Type A	M&F	12–17.9	O
Stevens & Truss (1985)	85	Edwards Personal Preference Schedule	E, A, C, O	M&F	18–21.9	S
L. H. Stewart (1964)	89	Omnibus Trait Inventory	E, C, N, O	M&F	18–21.9	S
Stricker & Ross (1964)	41	Myers-Briggs Type Indicator	E, A, C, O	M&F	18–21.9	S
Thomas & Chess (1986)	131	New York Longitudinal Study Temperament Scales	Act, NA, Pers, Ad, App, Rhy, Thresh	M&F	0–2.9 3–5.9	O
Tomlinson, Harbaugh, & Anderson (1996)	40	a. Infant Behavior Questionnaire b. Children's Behavior Questionnaire	Act, Pers, Ad	M&F	0–2.9	O
Torgersen (1988)	86	New York Longitudinal Study Temperament Scales	Act, NA, Pers, Ad, App, Rhy, Thresh	M&F	0–2.9 6–11.9	O
Troy (1988)	96	California Child Q-Sort	C, N	M&F	3–5.9 6–11.9	O
Tubman, Lerner, Lerner, & Von Eye (1992)	129	New York Longitudinal Study Temperament Scales	Ad	M&F	12–17.9 18–21.9	S
Tubman & Windle (1995)	975	Dimensions of Temperament Survey—Revised	Ad	M&F	12–17.9	S
Tuddenham (1959)	72	Observer ratings	E, A, C, N, O	M&F	12–17.9 18–21.9	O
Usala & Hertzog (1991)	227	Jackson Personality Inventory Trait Anxiety Scale	N	M&F	50–59	S
Viken, Rose, Kaprio, & Koskenvuo (1994)	4,922	Esyenck Personality Inventory	E, N	M&F	18–21.9 22–29 30–39 40–49 50–59	S
Von Dras & Siegler (1997)	3,318	a. Minnesota Multiphasic Personality Inventory b. NEO Personality Inventory	E	M&F	18–21.9	S
Weiss (1980)[d]	121	Myers-Briggs Type Indicator	E, A, C, O	M&F	18–21.9	S
Westenberg & Gjerde (1999)	98	Washington University Sentence Completion Test	O	M&F	12–17.9	P
Wheeler & Schwartz (1989)	225	Millon Clinical Multiaxial Inventory	E, A, C, N	M&F	18–21.9	S

(*table continues*)

Table 1 (*continued*)

Authors	*N*	Measures	Type of trait	Gender	Age category	Method
Whitbourne, Zuschlag, Elliot, & Waterman (1992)	238	Inventory of Psychosocial Development	A, C, N, O	M&F	18–21.9 30–39	S
Wiggins & Winder (1961)	339	Peer Nomination Inventory	E, A, N	M	6–11.9	O
Wilhelm & Parker (1990)	163	a. Eysenck Personality Inventory b. Depressive Experiences Questionnaire c. Costello-Comrey Trait Depression d. Bem Sex Role Inventory	A, N, F/M	M&F	22–29	S
Woodall & Matthews (1993)	85	Minnesota Multiphasic Personality Inventory	A	M&F	12–17.9	S
Woodruff (1983)	77	California Test of Personality	N	M&F	18–21.9	S
Yonge & Regan (1975)	833	Omnibus Personality Inventory	E, A, C, N, O, F/M	M	18–21.9	S

Note. Age category represents age at initiation of wave of longitudinal assessment. Type of trait: E = extraversion; A = agreeableness; C = conscientiousness; N = neuroticism; O = openness to experience; F/M = femininity/masculinity; Act = activity level; NA = negative affect; Pers = task persistence; Ad = adaptability; App = approach/withdrawal; Rhy = rhythmicity; Thresh = threshold. Gender: F = female; M = male; M&F = male and female. Method: S = self-report; O = observer rated; P = projective.

[a] Data from unpublished conference paper reproduced in Guilford, Zimmerman, and Guilford (1976). [b] Raw data on test–retest reliabilities of medical students at St. Mary's Hospital Medical School, cited in Myers and McCaulley (1993). [c] Raw data on retest reliability of the Myers-Briggs Type Indicator, cited in Myers and McCaulley (1993). [d] Longitudinal data of the University of New Mexico Nursing Program, cited in Myers and McCaulley (1993).

Table 2 presents information about the studies and coefficients culled from each study before the database was aggregated to analyze the effects of moderator variables. The average study in the initial database had a time interval of approximately 6.7 years, studied college students (mean age = 17.84), and had an attrition rate of 42%. Slightly more coefficients were based on reports from male participants than from female participants, and a large proportion of the studies did not analyze trait consistency separately for male and female participants. Fifty percent of coefficients were derived from self-reports, 41% from observer ratings, and 9% from projective tests. The majority of coefficients (77%) were coded into one of the Big Five categories, followed by temperament (20%), then non-Big-Five categories (3%).

The Relation Between Age and Trait Consistency

We expected trait consistency to steadily increase through childhood, adolescence, and early adulthood, then to plateau sometime after age 20. Figure 1 shows the estimated population test–retest correlations and 95% confidence-level estimates for each age category. In addition, Table 3 shows the above information and heterogeneity estimates of trait consistency across 10 age categories from infancy to age 73.[2] Two points of caution should be stated before interpreting these results. First, some studies contributed to population estimates for several age periods. Therefore, it would have been inappropriate to perform statistical tests comparing different age categories unless they did not contain information from overlapping samples. Second, we found significant heterogeneity in the estimated population correlation coefficients in all age categories. The significant heterogeneity estimates indicate that the estimated mean population correlation coefficients may vary significantly depending on numerous potential moderators of consistency.

Overall, trait consistency increased in a linear, steplike pattern until the ages 50 to 59 decade, when it peaked. Consistent with the argument that the earliest years of life may be too unstable to support traitlike patterns, trait consistency in the ages 0 to 2.9 period was the lowest at .35. Trait consistency increased dramatically to .52 in the 3 to 5.9 age period.[3] Between the ages 3 to 5.9 period and the college years, trait consistency essentially leveled off, with a slight drop between the second and third age periods and

Table 2
Descriptive Statistics for Characteristics Associated With Trait Consistency

Study characteristics	*M*/%	*SD*	Range
Time interval in years	6.75	7.51	1 to 53 years
Age in years at initiation of longitudinal study	17.84	15.72	6 weeks to 73 years
Attrition	42%	.24	0% to 96%
Male samples	34%		
Female samples	30%		
Both	36%		
Observer ratings	41%		
Self-reports	50%		
Projective tests	9%		
Big Five traits	77%		
Non-Big-Five traits	3%		
Temperament traits	20%		

Note. Descriptive statistics are based on overall database before aggregation.

[2] The sample sizes used to weight the estimates of trait consistency were by necessity estimates also. Mostly because of attrition, sample size often changed across multiple waves of assessment within a single longitudinal study. When multiple waves of data were averaged, the average sample size across assessment waves was used to create the weighting variable, confidence intervals, and heterogeneity estimates.

[3] One explanation for the dramatic increase in consistency in the ages 3 to 5.9 period is the relation between sample size and effect size found in that age category. For the ages 3 to 5.9 period, as well as several other age periods, the largest effect size corresponded to the largest sample size. When estimates were then weighted by the inverse of the variance (e.g., $N - 3$), the larger effect sizes contributed disproportionately to the population effect size. The unweighted effect sizes show a more linear trend (e.g., .35, .44, .46, .39, .54, .52, .49, .59, .74, and .62, respectively, for each age category shown in Table 3).

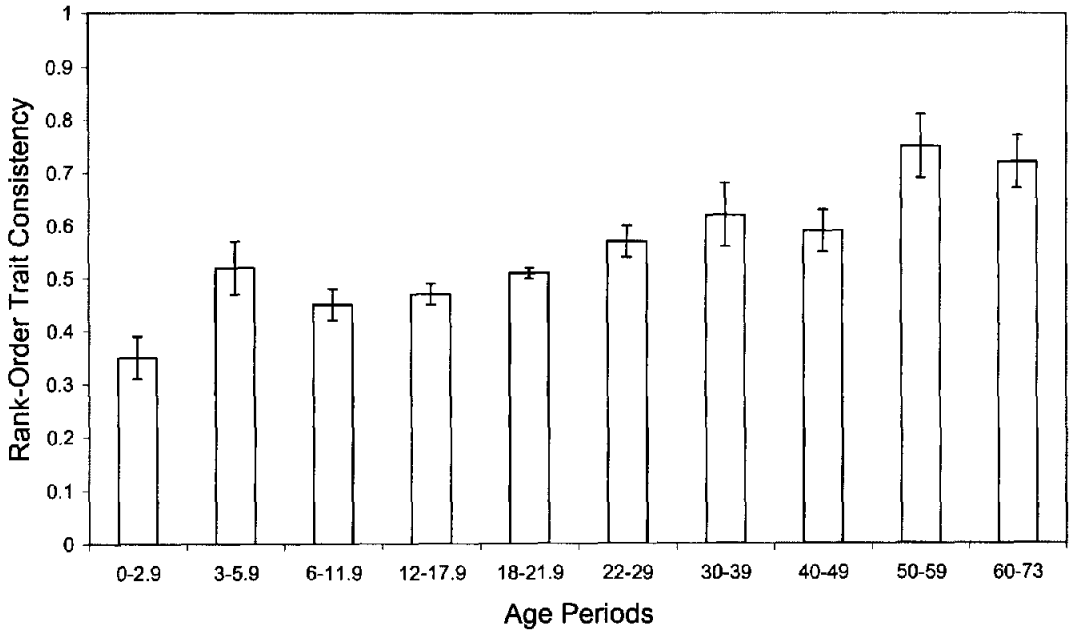

Figure 1. Population estimates of mean consistency across age categories (in years) with 95% confidence level estimates.

then a gradual increase until the college years (18–21.9). Between the college years and the first decade of young adulthood (ages 22–29), trait consistency again increased substantially from .51 to .57, where it once again leveled off. Trait consistency increased dramatically one more time between the first half of middle age (40–49) and the second half of middle age (50–59). Once again, trait consistency reached a plateau, as the estimate from the ages 60 to 73 period was quite similar to the ages 50 to 59 estimate.

Examination of the confidence-level estimates supports the notion that trait consistency increased at three points in the life course: from infancy and toddlerhood to the preschool period, from the college years to the early stages of young adulthood, and from early middle age to later middle age. For each of these transitions, the respective estimates of rank-order consistency from the earlier age periods fell below the 95% confidence interval (CI) estimate of the next age period. Within periods of the life course when trait consistency appears to plateau, all of the respective trait consistency estimates fell within each 95% CI estimate. In respect to our first question, it is apparent that the relation between age and trait consistency is linear and positive. That is, trait consistency increases with age. In respect to our second question concerning when trait consistency peaks, it appears that the peak occurs sometime after age 50. Our third question was whether trait consistency would peak near unity. The peak level of consistency after age 50 was well below unity. The latter finding would thus support the contention that personality traits do not stop changing at a specific point in the life course.

Table 3
Population Estimates of Trait Consistency Across Age Categories

Age (years)	ρ	K	CI	Q	N	ρ_t
0–2.9	.35	18	.31–.39	40.88*	2,085	.31
3–5.9	.52	12	.47–.57	67.14*	1,489	.49
6–11.9	.45	29	.42–.48	111.22*	4,053	.43
12–17.9	.47	32	.46–.48	153.85*	10,951	.43
18–21.9	.51	45	.50–.52	168.15*	11,340	.54
22–29	.57	10	.54–.60	59.91*	3,394	.60
30–39	.62	8	.56–.68	107.72*	1,055	.64
40–49	.59	11	.55–.63	55.42*	2,711	.60
50–59	.75	4	.69–.81	53.57*	948	.74
60–73	.72	6	.67–.77	78.20*	1,385	.71

Note. ρ = estimated population correlation; K = number of samples; CI = 95% confidence interval of estimated population correlation; Q = heterogeneity statistic; ρ_t = estimated population correlation controlling for time interval of longitudinal study.
* $p < .05$.

The Relation of Time Interval and Attrition to Trait Consistency

These initial population estimates were made without controlling time interval or attrition.[4] We used the first age-aggregated

[4] One approach to meta-analysis is to account for study artifacts, such as attrition and range restriction (see, e.g., J. E. Hunter & Schmidt, 1990) before estimating population averages. Because the focus of our study was, in part, on one of those artifacts (e.g., test–retest reliability), we were more interested in the relations among some of these indicators than in accounting for each artifact before estimating trait consistency. One common artifact that is controlled for is unreliability. In the case of test–retest correlation coefficients, the accepted norm is to correct for unreliability using short-term test–retest estimates of the same measure drawn from the same sample (Heise, 1969). The use of this indicator of reliability in the present study would have precluded using much of the data. We did compile interrater reliability estimates and internal consistency reliability estimates for each measure when possible. The mean estimates of these indices of reliability demonstrated little relation with age category—0–2.9 = .77, 3–5.9 = .76, 6–11.9 = .79, 12–17.9 = .74, 18–21.9 = .79,

database to test the relation of time interval and attrition to trait consistency. Consistent with previous research, the relation between time interval and consistency was negative and of modest size ($r = -.20$, $p < .05$). When the trait consistency estimates were weighted by the inverse of the variance and age was controlled for in a hierarchical regression equation, the standardized beta weight was larger at $-.36$, $p < .05$. Using the unstandardized beta weight estimates from the regression equation including age and time as predictors of trait consistency, we estimated the average trait consistency one could expect for various lengths of time. For these estimates, we held age constant at 20. On the basis of the present data, the average trait consistency over a 1-year period would be .55; at 5 years, it would be .52; at 10 years, it would be .49; at 20 years, it would be .41; and at 40 years, it would be .25.

Because the estimates of trait consistency may have been affected by time interval, we reestimated the population test–retest correlation coefficients depicted in Table 3. To compute the reestimated coefficients, we used an analysis of covariance (ANCOVA) model, with time interval as a covariate. This is the same as estimating the test–retest estimates as if all studies lasted the average interval of 6.7 years. The new population estimates are also shown on Table 3. When we controlled for time interval, the estimates in young adulthood were slightly lower than the original estimates. Overall, though, the effect of controlling for time interval was small, averaging about .02 across all age categories.

We also tested the assumption that increased attrition would lead to higher consistency. On the basis of the data from the age-aggregated database, the unweighted relation between attrition and trait consistency was small ($r = .06$, $p > .05$). We then reestimated the relation between attrition and trait consistency, holding constant both age and time interval and weighting the trait consistency estimates by the inverse of the variance. The effect of attrition remained quite small ($\beta = -.04$, $p > .05$).

The Relation of Gender and Method to Trait Consistency

Our next goal was to examine the relations of gender and method to trait consistency. Unfortunately, these two variables were not well represented across all age categories. Rather than examining the effect of each of these variables within each age category, we decided to collapse across age categories and control for the effect of time and age using ANCOVA estimates of trait consistency.

As described in the *Procedure* section, we created aggregated databases to examine the potential moderator effects of gender and method. One overall finding should again be noted before interpreting differences across gender and method categories. All of the estimates obtained significant heterogeneity except those for projective tests (see Table 4). Thus, it would be premature to consider these as invariant population estimates of trait consistency across the two moderator variables.

We first examined the average consistency for male, female, and mixed samples (see Table 4). Studies that did not break down results by gender obtained the highest population estimate of trait consistency at .51. Longitudinal studies of men and women obtained the same level of trait consistency at .49. After we controlled for age and time interval, the mixed gender samples again had the highest estimates of trait consistency ($M = .52$), followed by men at .49 and women at .48. The lack of gender differences supports previous research showing little or no difference between men and women on trait consistency (Schuerger et al., 1989).

Table 4
Population Estimates of Trait Consistency for Gender and Method

Gender and method	ρ	K	CI	Q	N	ρ_{ta}
Men	.49	48	.47–.51	220.28*	10,254	.49
Women	.49	43	.47–.51	224.32*	9,682	.48
Mixed	.51	72	.49–.53	703.70*	17,825	.52
Observer	.48	54	.46–.50	275.08*	7,594	.51
Self-report	.52	73	.51–.53	672.45*	22,908	.50
Projective	.43	8	.34–.52	9.33	489	.45

Note. ρ = estimated population correlation; K = number of samples; CI = 95% confidence interval of estimated population correlation; Q = heterogeneity statistic; ρ_{ta} = estimated population correlation controlling for time interval of longitudinal study and age of sample.
* $p < .05$.

We next examined mean trait consistency for different methods of assessment (see Table 4). Self-report, observer, and projective methods were distributed quite unevenly across different ages. Self-reports were rare in the first decade of life, whereas observer methods were rare in adulthood. Projective tests were confined mostly to high school and college age samples. Despite the uneven distribution of methods across age categories, the initial estimates for the different types of method were similar in magnitude. Without controlling for age and interval, self-report methods proved to be the most consistent at .52, followed by observer methods (.48), and projective tests (.43). Once age and interval were controlled, observer and self-report methods reached similar levels of consistency (.51 and .50, respectively), followed by projective tests at .45. According to the confidence-level estimates, the mean self-report consistency when adjusting for sampling error only was higher than observer or projective methods. Given the small magnitude of the difference, we feel the most impressive feature of these analyses is the lack of substantive differences between these three primary methods of assessing traits.

The Effect of Type of Trait on Trait Consistency

In our final set of analyses, we tested the moderating effect of type of trait on trait consistency. Temperaments and traits were almost categorically distributed across age. That is, all temperament estimates were derived from age periods before college with 99% of the estimates coming before age 12. Conversely, no studies reported assessing adult personality traits before age 3. Adult personality traits became a more common source of data in the middle childhood period (ages 6–11.9) and beyond. Therefore, in addition adjusting the analyses for sampling error, we also estimated trait consistency for temperament and adult trait categories holding constant age and time interval.

22–29 = .81, 30–39 = .69, 40–49 = .75, 50–59 = .78, 60–73 = .78—and were relatively unrelated to the trait consistency estimates (uncorrected $r = .13$).

Previous research did not test whether traits from different temperament categories exhibit differential consistency, so we had no hypotheses for the temperament categories. For the estimates adjusted for sampling error alone, the highest consistency was exhibited by the scales categorized as measuring adaptability (see Table 5). Of the remaining temperament dimensions, approach, negative emotionality, task persistence, and rhythmicity all obtained similar levels of consistency ranging from .35 to .41. Activity level and threshold showed the lowest levels of consistency. All of the trait consistency estimates demonstrated heterogeneity of variance except for scales categorized in the threshold domain according to the Q statistics. The lack of heterogeneity for the threshold category is most likely a result of the small number of longitudinal samples in which threshold measures were tracked. Table 5 also shows the estimates of consistency for temperament when age and time interval were held constant. As would be expected from the strong linear effect of age on trait consistency and the fact that most temperament measures were drawn from early childhood, consistency of temperaments increased when age and time interval were controlled.

As one would expect, the consistency estimates for the adult personality traits were uniformly higher than those for the temperament traits (see Table 5). Within the adult personality trait categories, previous research on trait consistency found that scales assessing extraversion were the most consistent, but these studies did not assess the full spectrum of the Big Five. Our findings show that measures of extraversion and agreeableness were the most consistent (M = .55 for both trait categories). The remaining Big Five measures were also quite consistent, ranging from .50 to .52. According to the CIs, traits drawn from the extraversion and agreeableness domains were more consistent than traits drawn from any of the remaining trait categories, although the magnitude of the difference is quite small. Of the non-Big-Five trait categories, both the femininity/masculinity and Type A scales were as consistent as most of the Big Five traits. When we controlled age and time interval in an ANCOVA, the estimates remained essentially unchanged except for estimates of the consistency of Type A measures, which dropped.

Table 5
Population Estimates of Trait Consistency for Temperaments and Traits

Categories	ρ	K	CI	Q	N	ρ_{ta}
Temperament						
Approach	.41	12	.36–.46	44.47*	1,625	.51
Adaptability	.47	14	.43–.51	60.26*	2,465	.52
Task persistence	.36	14	.31–.41	75.32*	1,667	.47
Negative emotionality	.35	9	.39–.51	17.28*	1,137	.46
Activity level	.28	9	.22–.36	18.13*	981	.41
Rhythmicity	.39	5	.32–.46	22.38*	893	.49
Threshold	.21	4	.12–.30	5.37	446	.35
Adult personality trait						
Extraversion	.54	67	.53–.55	714.10*	20,711	.55
Agreeableness	.54	47	.52–.56	376.70*	8,428	.52
Conscientiousness	.51	51	.49–.53	423.59*	11,513	.49
Neuroticism	.50	68	.48–.52	711.49*	15,118	.46
Openness	.51	50	.49–.53	239.52*	7,901	.51
Femininity/masculinity	.52	17	.49–.55	93.65*	4,340	.51
Type A	.52	4	.44–.60	21.65*	634	.41

Note. ρ = estimated population correlation; K = number of samples; CI = 95% confidence interval of estimated population correlation; Q = heterogeneity statistic; ρ_{ta} = estimated population correlation controlling for time interval of longitudinal study and age of sample.
* $p < .05$.

Discussion

Over the last 30 years, the field of personality psychology has witnessed a dramatic swing in opinion regarding the scientific merits of traits. Opinions have shifted from the perspective that traits are ephemeral concepts (see, e.g., Mischel, 1968, 1992) and that situations are largely responsible for behavior (Ross & Nisbett, 1991) to the perspective that personality traits are so stable they are essentially fixed in adulthood (McCrae & Costa, 1994). The first and possibly most important finding of the present meta-analysis is the relatively high levels of consistency demonstrated across the life course. These estimates easily exceed the unfortunate .30 figure promulgated by critics of the trait construct (see, e.g., Mischel, 1968). In turn, the trait consistency estimates are not so high as to warrant the conclusion that no change occurs in adulthood. It appears then that traits are mostly consistent in adulthood, with some indication that they retain a dynamic quality (Pervin, 1994), a moderate position acknowledged in the past but often ignored (see, e.g., Block, 1971; Helson & Wink, 1992; Kogan, 1990; Olweus, 1979).

Age and Trait Consistency

Understanding the relation between age and trait consistency is critical to fields such as personality, developmental, clinical, and industrial psychology. Despite its importance, the topic of trait consistency has been primarily the focus of individual longitudinal research teams (e.g., Haan, Millsap, & Hartka, 1986; Helson & Wink, 1992) or narrative reviewers of longitudinal research (see, e.g., Aldwin & Levenson, 1994; Costa & McCrae, 1997). Evidence for the relation between age and trait consistency has not been compiled quantitatively across many longitudinal studies since Bloom (1964). Bloom concluded that the data did not support the psychoanalytic theories placing the final development of traits in childhood or late adolescence. However, at the time of Bloom's review, the longitudinal database lacked a sufficient number of studies of adults to draw conclusions concerning the consistency of traits beyond age 20.

Since Bloom's (1964) review, numerous longitudinal studies of temperament and personality traits from childhood through old age have been published. Our quantitative review of these studies shows that trait consistency increases in a linear yet steplike fashion from infancy to middle age where it then reaches its peak sometime after age 50. Consistent with the perspective that the earliest years of life are marked by the least consistency, the lowest estimate of trait consistency was found in the infant and toddler age period. Beyond the earliest years of life, trait consistency increases in a steplike function with increases coming in the preschool years, in young adulthood, and then again in middle age. After middle age, trait consistency reaches a plateau around .75.

Three hypotheses about the relation between age and trait consistency have been made. First, it is generally accepted that, with age, traits become increasingly consistent. This primary assumption of personality development is supported by our data, although the steplike pattern of increase was unexpected. Given the heter-

ogeneous nature of the data collected for this study (e.g., different methods, assessment systems) and the modest number of samples in any one age category (4 to 45), we are reticent to identify these "steps" as key transition points in the life course without further investigation. For example, if one sets aside the relatively high estimate in the preschool period, trait consistency shows a strong linear increase from infancy to the age 30 to 39 decade. Furthermore, even though some of the life course patterns of the factors associated with consistency are identified, such as environmental and identity patterns, explanations for the increases over specific periods are lacking.

The second hypothesis was that trait consistency should peak sometime in adulthood. Bloom (1964) and McCrae and Costa (1994) argued that personality traits should peak at age 20 or 30, respectively. These arguments are not supported by the meta-analytic estimates. Trait consistency did not peak until after age 50. Alternatively, we proposed, based on a variety of theories and empirical findings, that trait consistency would peak in middle age. The results of our review support this hypothesis best. The middle age hypothesis was based on the knowledge that factors presumed to be associated with higher estimates of trait consistency, such as identity certainty, become more prominent in middle age (see, e.g., A. J. Stewart & Ostrove, 1998). Achieving a strong identity is conceptually linked to many of the other factors associated with increased consistency, such as the ability to choose environments that fit well with one's identity, the propensity to evoke consistency-engendering responses from others, and the ability to assimilate more experience. Likewise, middle age was identified by Neugarten (1968) as the time of the executive personality, which is characterized by increased mastery, control over the environment, and the ability to better cope with life's complexities. It is clear that our findings support the theory that trait consistency peaks in middle age. What remains to be tested is the extent to which all of the factors hypothesized to be linked to increased consistency, such as a stable environment, identity integration, and stable person–environment transactions, are themselves associated with age and increasing trait consistency.

Finally, some researchers have argued that personality traits actually stop changing in adulthood. More specifically, some have claimed that personality stops changing at age 30 (McCrae & Costa, 1994). Therefore, we should not only have found that trait consistency peaks, but that it peaks close to unity. Obviously, the age 30 estimate was inaccurate, but there remains the possibility that the peak after age 50 is high enough to support the conclusion that personality traits are essentially fixed at this age. Before drawing this conclusion, we would again point out that the construct of personality change includes at least four indicators (rank-order, mean-level, ipsative, and individual differences) and that conclusions concerning the changeability of traits would profit from investigating all four. For example, at least one study has shown the existence of mean-level change in personality traits well beyond age 50 (see, e.g., Field & Millsap, 1991). Moreover, each of these indicators of change is relatively unrelated to the other (Block, 1971; M. L. Kohn, 1980; Ozer, 1986), which means that perfect unity in rank-order consistency would be insufficient evidence to rule out change in any or all of the three remaining indexes of consistency.

One factor that could affect the relation between trait consistency and age that has not been considered in previous theories is the historical context in which the longitudinal studies were embedded (see, e.g., Elder, 1979). That is, most of the theories and data on personality trait consistency come from the late 20th century. Ironically, McCrae and Costa's (1994) perspective that traits are fixed by age 30, based on James (1890/1950), may have been correct for people living at the end of the 19th century. Before 1920, less than 16% of the populace completed high school, and most people left school by age 16 to start a career (Modell, 1989). Thus, many people were entering their careers in their late teens and early twenties and by age 30 would have been in their careers for 14 years and most likely married and with children (Modell, 1989). Furthermore, life expectancy at this time was approximately 55. This combination of life context factors and life expectancy limitations during this period in history would mean that age 30 may have corresponded to middle age.

In contrast, the generations that followed increasingly acquired more schooling, delayed their careers, and delayed their development of a strong identity (Littwin, 1986). With the increasing effectiveness of public health interventions, the life span steadily increased. Current generations now confront a life course in which childhood and adolescence can stretch into one's twenties, marriage and children can be delayed well into one's thirties, and retirement can be put off indefinitely (Modell, 1989). One plausible explanation of our findings is that, since the turn of the century, consecutive generations have stretched the trajectory of the life course and at the same time stretched the time it takes to fully develop one's traits.

We are left with the possibility that, in the present historical context, life provides continuing challenges to the consistency of personality well into adulthood. As can be concluded from the less than perfect rank-order consistency after age 50, there may remain unforeseen experiences and demands, such as retirement, that potentially influence trait consistency well into old age (see Field & Millsap, 1991).

The Effect of Time, Attrition, Gender, and Method on Trait Consistency

As has been repeatedly shown in previous research, the longer a longitudinal study lasts, the lower are its estimates of trait consistency (see, e.g., Conley, 1984a; Crook, 1941; Schuerger et al., 1989). It is important to understand the effect of time on the estimates depicted in the results. For example, if we had studied short-term longitudinal studies, our estimates would have been several points higher. Conversely, if we had studied longer longitudinal studies, our estimates would have been lower. Interestingly, controlling for time interval did not dramatically change the estimates of trait consistency. This was most likely because the aggregation of multiple studies within each age category essentially controlled for time interval.

It is common to criticize longitudinal studies that exhibit excessive levels of attrition because of the unsystematic nature of the resulting sample. Furthermore, it is often assumed that attrition adversely affects the results of longitudinal investigations of trait consistency. For example, Finn (1986) argued that the high estimates of trait consistency reported by Costa, McCrae, and Arenberg (1980) were the result of high attrition levels. The logic behind this claim is that the participants who remain in a longitudinal study are, by their continued participation, showing evidence of being more consistent. Thus, estimates of trait consistency drawn from studies with high attrition could be inflated because of

biases in the resulting sample. In contrast to these expectations, across hundreds of longitudinal studies, we did not find that attrition distorted the resulting trait consistency. Obviously, this finding should not lead to the conclusion that attrition be ignored in longitudinal research. For example, many longitudinal studies rely on samples drawn from privileged and educated populations. Therefore, an existing longitudinal database may be biased toward a population of individuals that is more consistent than the norm, diminishing the effect of attrition. However, this finding does call into question the untested assumption that attrition is necessarily a distorting influence and warrants a more systematic investigation of the effects of attrition on longitudinal study results.

Several study and scale factors had little or no effect on rank-order consistency. As in previous research (e.g., Schuerger et al., 1989), studies focusing on men or women did not show systematic differences in trait consistency over time. Method of assessment proved to be confounded with the age of the population under study. Most observer methods were used with children. Most self-report methods were used with adults, and projective tests were mostly concentrated on high school and college age samples. Regardless of the uneven distribution of methods across the life course, the differences in trait consistency among these three methods were quite small. This lack of differences among methods could be taken as an indication of the robust nature of trait consistency. We would be reticent to infer from this analysis that all methods have equal merit, as we did not test the validity of each technique. Furthermore, without a better representation across different ages, it would be appropriate to defer a conclusion until more studies using a variety of methods at different ages are completed.

The Effect of Type of Temperament and Trait on Trait Consistency

The assumption of longitudinal consistency has been a strong feature of the definition of temperament since the concept was introduced (see, e.g., A. H. Buss & Plomin, 1975). In the present study, the estimates for the consistency of temperament constructs ranged from low to moderate with most in the moderate range. Several features of the studies reviewed could be used to argue that these are underestimates of temperament consistency. First, the studies compiled in the present database were all 1 year or longer in duration, therefore ignoring the potentially higher levels of consistency that may be found in shorter longitudinal studies of temperament. Shorter longitudinal studies may be more appropriate for the tracking of consistency of temperament, which is assumed to fluctuate substantially over the initial years of the life course. Second, most longitudinal studies of temperament employed different questionnaires and rating systems at different ages. That is, developmental researchers have often created age-specific measures of temperament constructs that are conceptually similar but often use different items to rate each temperament dimension. The use of different measures at different ages may contribute to an underestimate of temperament consistency. In summary, most temperament categories demonstrated moderate levels of consistency that could possibly be higher if certain methodological factors could be addressed in future research.

Four of the five most well-accepted dimensions of temperament (approach, negative emotionality, task persistence, and adaptability) showed moderate levels of consistency. Interestingly, the fifth domain—activity level—showed less consistency. R. P. Martin et al. (1994) offered several insights as to why activity level may demonstrate lower consistency. First, activity level is difficult to distinguish from emotional reactivity in early infancy. Second, as children move into social contexts that limit activity, such as elementary school, the opportunity to observe activity level is diminished, as is the ability to differentiate children on the construct. Conversely, activity level may become part of qualitatively different traits as children age. For example, Digman and Inouye (1986) suggested that descriptions of activity were related to extraversion (positively) and conscientiousness (negatively), indicating that as children age, they may find an outlet for activity level in social interactions and impulsivity (see also Eaton, 1994). Furthermore, Graziano et al. (1988) showed that activity level in childhood was linked to caregiver expectations for children to become extraverted as adults. Therefore, activity level may be transformed into adult personality partially through the expectations of others.

The existence of the remaining two categories of temperament, rhythmicity and threshold, is controversial because of potential measurement confounds (e.g., the widespread use of the New York Longitudinal Study temperament measurement system). Regardless of its questionable construct validity, the rhythmicity domain exhibited levels of consistency equal to the five accepted temperament constructs. The relatively high consistency of the rhythmicity measures may warrant further investigation of this construct. In contrast, the threshold category demonstrated the lowest consistency of all temperament categories. This low consistency may be attributable, in part, to the low number of studies that followed threshold over time.

As one would expect from the findings on the effect of age on consistency, adult measures of traits were more consistent than measures of temperament. Previous studies have reported that measures of extraversion were more consistent than measures of neuroticism (Conley, 1984a; Schuerger et al., 1989). Although we replicated this finding in the present study, the general conclusion that the domain of extraversion is the most consistent trait domain was not supported. Previous research did not include an examination of trait consistency across all of the Big Five dimensions, such as agreeableness, conscientiousness, and openness to experience. Interestingly, we found that measures of extraversion and agreeableness were the most consistent Big Five traits. One should keep in mind, however, that the means for extraversion and agreeableness exceeded those of other traits by only approximately 3 to 4 hundredths of a point.

Conclusion

The findings of our meta-analysis must be placed in the context of our definition of consistency. We focused exclusively on rank-order consistency. It would be inappropriate to draw inferences from these data about other statistical approaches to describing consistency (e.g., mean-level, ipsative, and absolute consistency). Any attempt to draw conclusions concerning the changeability or consistency of personality traits in general would by necessity have to follow the same prescription. Without accounting for the full range of approaches to consistency, it would be premature to render a final judgment concerning whether and when personality traits are fixed. Furthermore, several holes in the research literature have yet to be filled. For example, we found little information regarding personality consistency in the later years of the life span. Additional studies of person-

ality that focus on the elderly will provide invaluable insights as to whether personality consistency continues to increase, decline, or remain at the levels found in middle age.

Another limitation to the present review was the necessary use of chronological age in estimating the relation between consistency and age. Chronological age is only one of several ways to estimate age and may not be the most relevant indicator of trait consistency (Birren & Cunningham, 1985). For example, Birren and Cunningham (1985) have also described social age, which refers to the timing of a person's roles and habits, and psychological age, which reflects the behavioral capacities of individuals. Both of these alternative indicators of age may be more relevant for personality consistency than is chronological age. Support for these alternative types of age having differential influence over personality processes comes from the research on achievement patterns of eminent people (Simonton, 1977). Simonton (1977) found that people who showed early achievement died sooner than people who showed achievement later in life. Thus, people who show early achievement in life may have an accelerated life course both socially and psychologically. Future research on the relation between age and trait consistency may profit from assessing alternative indicators of age, such as social or psychological age.

In conclusion, the results from our meta-analysis support the inference that traits are quite consistent over the life course. The results do not support the hypothesis that traits reach a plateau early in the life course. Rather, according to rank-order estimates of personality traits, consistency peaks after age 50 at a level not high enough to infer a complete lack of change in personality traits.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Adams, G. R., & Fitch, S. A. (1981). Ego stage and identity status development: A cross-lag analysis. *Journal of Adolescence, 4,* 163–171.

*Adams, S. H. (1994). Role of hostility in women's health during midlife: A longitudinal study. *Health Psychology, 13,* 488–495.

Ahadi, S. A., & Rothbart, M. K. (1994). Temperament, development, and the Big Five. In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 189–207). Hillsdale, NJ: Erlbaum.

Alder, A. G., & Scher, S. J. (1994). Using growth curve analyses to assess personality change and stability in adulthood. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 149–174). Washington, DC: American Psychological Association.

Aldwin, C. M., & Levenson, M. R. (1994). Aging and personality assessment. In M. P. Lawton & J. A. Teresi (Eds.), *Annual review of gerontology and geriatrics: Focus on assessment* (pp. 182–209). New York: Springer.

*Asendorpf, J. B. (1990). Development of inhibition during childhood: Evidence for situational specificity and a two-factor model. *Developmental Psychology, 26,* 721–730.

*Asendorpf, J. B., & van Aken, M. A. G. (1991). Correlates of the temporal consistency of personality patterns in childhood. *Journal of Personality, 59,* 689–703.

*Backteman, G., & Magnusson, D. (1981). Longitudinal stability of personality characteristics. *Journal of Personality, 49,* 148–160.

*Baltes, P. B., & Nesselroade, J. R. (1972). Cultural change and adolescent personality development: An application of longitudinal sequences. *Developmental Psychology, 7,* 244–256.

*Bar-Tal, D., & Raviv, A. (1979). Consistency of helping-behavior measures. *Child Development, 50,* 1235–1238.

*Bates, M. E., & Pandina, R. J. (1989). Individual differences in the stability of personality needs: Relations to stress and substance use during adolescence. *Personality and Individual Differences, 10,* 1151–1157.

Bell, R. Q., & Chapman, M. (1986). Child effects in studies using experimental or brief longitudinal approaches to socialization. *Developmental Psychology, 22,* 595–603.

Bem, S. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42,* 165–172.

Birren, J. E., & Cunningham, W. R. (1985). Research on the psychology of aging: Principles, concepts and theory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (2nd ed., pp. 3–34). New York: Van Nostrand Reinhold.

*Block, J. (1971). *Lives through time.* Berkeley, CA: Bancroft Books.

*Block, J. (1993). Studying personality the long way. In D. C. Funder, R. D. Parke, C. Tomlinson-Keasey, & K. Widaman (Eds.), *Studying lives through time* (pp. 9–41). Washington, DC: American Psychological Association.

*Block, J., Block, J. H., & Harrington, D. M. (1974). Some misgivings about the Matching Familiar Figures Test as a measure of reflection-impulsivity. *Developmental Psychology, 10,* 611–632.

Block, J., & Kremen, A. (1996). IQ and ego-resiliency: Clarifying their conceptual and empirical linkage and separateness. *Journal of Personality and Social Psychology, 70,* 349–361.

Block, J., & Robins, R. W. (1993). A longitudinal study of consistency and change in self-esteem from early adolescence to early adulthood. *Child Development, 64,* 909–923.

Bloom, B. S. (1964). *Stability and change in human characteristics.* New York: Wiley.

*Bolton, B. (1979). Longitudinal stability of the primary and secondary dimensions of the 16PF-E. *Multivariate Experimental Clinical Research, 4,* 67–71.

*Broberg, A., Lamb, M. E., & Hwang, P. (1990). Inhibition: Its stability and correlates in sixteen- to forty-month-old children. *Child Development, 61,* 1153–1163.

*Bromberger, J. T., & Matthews, K. A. (1996). A "feminine" model of vulnerability to depressive symptoms: A longitudinal investigation of middle-aged women. *Journal of Personality and Social Psychology, 70,* 591–598.

*Bronson, W. C. (1966). Central orientations: A study of behavior organization from childhood to adolescence. *Child Development, 37,* 125–155.

*Bullock, D., & Merrill, L. (1980). The impact of personal preference on consistency through time: The case of childhood aggression. *Child Development, 51,* 808–814.

Buss, A. H., & Plomin, R. (1975). *Temperament: Early developing personality traits.* Hillsdale, NJ: Erlbaum.

Buss, D. M. (1987). Selection, evocation, and manipulation. *Journal of Personality and Social Psychology, 53,* 1214–1221.

*Buss, D. M., Block, J. H., & Block, J. (1980). Preschool activity level: Personality correlates and developmental implications. *Child Development, 51,* 401–408.

*Cairns, E., McWhirter, L., Duffy, U., & Barry, R. (1990). The stability of self-concept in late adolescence: Gender and situational effects. *Personality and Individual Differences, 11,* 937–944.

Cairns, R. B., & Hood, K. E. (1983). Continuity in social development: A comparative perspective on individual difference prediction. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (pp. 301–358). New York: Academic Press.

*Cantoni, L. J. (1955). A study in emotional adjustment: The correlation of student and adult forms of the Bell Adjustment Inventory over a period of thirteen years. *Educational and Psychological Measurement, 15,* 137–143.

*Caputo, D. V., Psathas, G., & Plapp, J. M. (1966). Test–retest reliability of the EPPS. *Educational and Psychological Measurement, 26,* 883–886.

*Carmelli, D., Rosenman, R. H., & Chesney, M. A. (1987). Stability of the Type A Structured Interview and related questionnaires in a 10-year

follow-up of an adult cohort of twins. *Journal of Behavioral Medicine, 10,* 513–525.

*Carmichael, C. M., & McGue, M. (1994). A longitudinal study of personality change and stability. *Journal of Personality, 62,* 1–20.

Case, R., Hayward, S., Lewis, M., & Hurst, P. (1988). Toward a neo-Piagetian theory of cognitive and emotional development. *Developmental Review, 8,* 1–51.

Caspi, A. (1998). Personality development across the life course. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 311–388). New York: Wiley.

Caspi, A., Elder, G. H., Jr., & Bem, D. J. (1988). Moving away from the world: Life-course patterns of shy children. *Developmental Psychology, 24,* 824–831.

Caspi, A., & Herbener, E. S. (1990). Continuity and change: Assortative marriage and the consistency of personality in adulthood. *Journal of Personality and Social Psychology, 58,* 250–258.

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 psychology: Theory and research* (2nd ed., pp. 300–326). New York: Guilford Press.

Caspi, A., & Silva, P. A. (1995). Temperamental qualities at age 3 predict personality traits in young adulthood: Longitudinal evidence from a birth cohort. *Child Development, 66,* 486–498.

*Cattell, R. B., & Cattell, M. D. (1975). *Handbook for the Jr.-Sr. High School Personality Questionnaire "HSPQ."* Urbana, IL: Institute for Personality and Ability Testing.

Clarke, A. D. B., & Clarke, A. M. (1984). Constancy and change in the growth of human characteristics. *Journal of Child Psychology and Psychiatry, 25,* 191–210.

Clausen, J. A. (1993). *American lives: Looking back at the children of the Great Depression.* New York: Free Press.

Conley, J. J. (1984a). The hierarchy of consistency: A review and model of longitudinal findings on adult individual differences in intelligence, personality, and self-opinion. *Personality and Individual Differences, 5,* 11–26.

*Conley, J. J. (1984b). Longitudinal consistency of adult personality: Self-reported psychological characteristics across 45 years. *Journal of Personality and Social Psychology, 47,* 1325–1333.

*Conley, J. J. (1985). Longitudinal stability of personality traits: A multitrait-multimethod-multioccasion analysis. *Journal of Personality and Social Psychology, 49,* 1266–1282.

*Costa, P. T., Jr., & McCrae, R. R. (1977–1978). Age differences in personality structure revisited: Studies in validity, stability, and change. *International Journal of Aging and Human Development, 8,* 261–275.

*Costa, P. T., Jr., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology, 54,* 853–863.

*Costa, P. T., Jr., & McCrae, R. R. (1992). Multiple uses for longitudinal personality data. *European Journal of Personality, 6,* 85–102.

Costa, P. T., Jr., & McCrae, R. R. (1997). Longitudinal stability of adult personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 269–292). San Diego, CA: Academic Press.

*Costa, P. T., Jr., McCrae, R. R., & Arenberg, D. (1980). Enduring dispositions in adult males. *Journal of Personality and Social Psychology, 38,* 793–800.

Crick, N. R., & Dodge, K. A. (1994). A review and reformation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115,* 74–101.

*Crook, M. N. (1941). Retest correlations in neuroticism. *Journal of General Psychology, 24,* 173–182.

*Crook, M. N. (1943). A retest with the Thurstone Personality Schedule after six and one-half years. *Journal of General Psychology, 28,* 111–120.

*Davis, M. H., & Franzoi, S. L. (1991). Stability and change in adolescent self-consciousness and empathy. *Journal of Research in Personality, 25,* 70–87.

*Davis, T. N., & Satterly, D. J. (1969). Personality profiles of student teachers. *British Journal of Educational Psychology, 39,* 183–187.

*Denham, S. A., Lehman, E. B., Moser, M. H., & Reeves, S. L. (1995). Continuity and change in emotional components of infant temperament. *Child Study Journal, 25,* 289–308.

*Digman, J. M. (1989). Five robust trait dimensions: Development, stability, and utility. *Journal of Personality, 57,* 195–214.

Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology, 50,* 116–123.

Digman, J. M., & Shmelyov, A. G. (1996). The structure of temperament and personality in Russian children. *Journal of Personality and Social Psychology, 71,* 341–351.

*Dudek, S. Z., & Hall, W. B. (1991). Personality consistency: Eminent architects 25 years later. *Creativity Research Journal, 4,* 213–231.

*Dusek, J. B., & Flaherty, J. F. (1981). The development of the self-concept during adolescent years. *Monographs of the Society for Research in Child Development, 46,* (4, Serial No. 191).

Eaton, W. O. (1994). Temperament, development, and the five-factor model: Lessons from activity level. In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 173–188). Hillsdale, NJ: Erlbaum.

Eder, R. A., & Mangelsdorf, S. C. (1997). The emotional basis of early personality development: Implications for the emergent self-concept. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 209–240). San Diego, CA: Academic Press.

*Eisenberg, N., Shell, R., Pasternack, J., Lennon, R., Beller, R., & Mathy, R. M. (1987). Prosocial development in middle childhood: A longitudinal study. *Developmental Psychology, 23,* 712–718.

Elder, G. H., Jr. (1979). Historical change in life patterns and personality. In P. B. Baltes & O. G. Brim, Jr. (Eds.), *Life-span development and behavior* (Vol. 2, pp. 117–159). New York: Academic Press.

*Englert, A. M. (1993). Stability and change in adolescence and adulthood: Analyses of data from the Genetic Studies of Genius (Doctoral dissertation, University of California, Riverside, 1993). *Dissertation Abstracts International, 54* (07), 3888B.

*Eron, L. D., Huesmann, L. R., Lefkowitz, M. M., & Walder, L. O. (1972). Does television cause aggression? *American Psychologist, 27,* 253–263.

*Exner, J. E. (1980). But it's only an inkblot. *Journal of Personality Assessment, 44,* 563–577.

*Exner, J. E., Thomas, E. A., & Mason, B. (1985). Children's Rorschachs: Description and prediction. *Journal of Personality Assessment, 49,* 13–20.

*Farnsworth, P. R. (1938). A genetic study of the Bernreuter Personality Inventory. *Journal of Genetic Psychology, 52,* 3–13.

*Farrington, D. P. (1978). The family backgrounds of aggressive youths. In L. A. Hersov, M. Berger, & D. Shaffer (Eds.), *Aggression and antisocial behaviour in childhood and adolescence* (pp. 73–94). Oxford, England: Pergamon Press.

Feldman, R. S. (1997). *Development across the life span.* Upper Saddle River, NJ: Prentice Hall.

*Field, D., & Millsap, R. E. (1991). Personality in advanced old age: Continuity or change? *Journal of Gerontology: Psychological Sciences, 46,* 299–308.

*Finn, S. E. (1986). Stability of personality self-ratings over 30 years: Evidence for an age/cohort interaction. *Journal of Personality and Social Psychology, 50,* 813–818.

Fischer, K. W., & Silvern, L. (1985). Stages and individual differences in cognitive development. *Annual Review of Psychology, 36,* 613–648.

Fiske, S. T., & Taylor, S. (1991). *Social cognition.* New York: McGraw-Hill.

*Ge, X., Lorenz, F. O., Conger, R. D., Elder, G. H., Jr., & Simons, R. L.

(1994). Trajectories of stressful life events and depressive symptoms during adolescence. *Developmental Psychology, 30,* 467–483.

*Gest, S. D. (1997). Behavioral inhibition: Stability and associations with adaption from childhood to early adulthood. *Journal of Personality and Social Psychology, 72,* 467–475.

*Giuganino, B. M., & Hindley, C. B. (1982). Stability of individual differences in personality characteristics from 3 to 15 years. *Personality and Individual Differences, 3,* 287–301.

Glenn, N. D. (1980). Values, attitudes, and beliefs. In O. G. Brim, Jr., & J. Kagan (Eds.), *Constancy and change in human development* (pp. 596–640). Cambridge, MA: Harvard University Press.

*Gold, S. R., & Henderson, B. B. (1990). Daydreaming and curiosity: Stability and change in gifted children and adolescents. *Adolescence, 25,* 701–708.

Goldberg, L. (1993). The structure of phenotypic personality traits. *American Psychologist, 48,* 26–34.

*Goldsmith, H. H. (1996). Studying temperament via construction of the Toddler Behavior Assessment Questionnaire. *Child Development, 67,* 218–235.

*Gough, H. G. (1987). *California Psychological Inventory: Administrator's guide.* Palo Alto, CA: Consulting Psychologists Press.

*Gough, H. G., & Bradley, P. (1996). *California Psychological Inventory Manual* (3rd Ed.). Palo Alto, CA: Consulting Psychologists Press.

Graziano, W. G., Jensen-Campbell, L. A., & Sullivan-Logan, G. M. (1998). Temperament, activity, and expectations for later personality development. *Journal of Personality and Social Psychology, 74,* 1266–1277.

*Grigoriadis, S., & Fekken, G. C. (1992). Person reliability on the Minnesota Multiphasic Personality Inventory. *Personality and Individual Differences, 13,* 491–500.

*Guerin, D., & Gottfried, A. W. (1987). Minnesota Child Development Inventories: Predictors of intelligence, achievement, and adaptability. *Journal of Pediatric Psychology, 12,* 595–609.

*Guerin, D., & Gottfried, A. W. (1994). Developmental stability and change in parent reports of temperament: A ten-year longitudinal investigation from infancy through preadolescence. *Merrill-Palmer Quarterly, 40,* 334–355.

Guilford, J. S., Zimmerman, W. S., & Guilford, J. P. (Eds.). (1976). *The Guilford-Zimmerman Temperament Survey handbook: Twenty-five years of research and application.* San Diego, CA: EdITS Publishers.

*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.

*Hagberg, B., Samuelsson, G., Lindberg, B., & Dehlin, O. (1991). Stability and change of personality in old age and its relation to survival. *Journal of Gerontology, 46,* 285–291.

*Hamlin, A. M. (1991). Adolescent personality, intervening life circumstances, and adult functioning (Doctoral dissertation, Cleveland State University, 1991). *Dissertation Abstracts International, 52* (07), 3939.

Harkness, A. R., & Hogan, R. (1995). The theory and measurement of traits: Two views. In J. N. Butcher (Ed.), *Clinical personality assessment: Practical approaches* (pp. 28–41). New York: Oxford University Press.

*Harkness, A. R., Spiro, A., III, Butcher, J. N., & Ben-Porath, Y. S. (1995, August). *Personality Psychopathology Five (PSY-5) in the Boston VA Normative Aging Study.* Paper presented at the 103rd Annual Convention of the American Psychological Association, New York.

*Harsany, M. (1993). The psychological functioning of elders entering old-old age: A longitudinal study (Doctoral dissertation, Concordia University, 1993). *Dissertation Abstracts International, 56* (07), 4033.

Harter, S. (1983). Developmental perspectives on the self-system. In P. H. Mussen (Series Ed.) & E. M. Hetherington (Vol. Ed.), *Handbook of Child Psychology: Vol. 4. Socialization, personality, and social development* (4th ed., pp. 275–385). New York: Wiley.

*Hathaway, S. R., & Monachesi, E. D. (1963). *Adolescent personality and behavior: MMPI patterns of normal, delinquent, dropout, other outcomes.* Minneapolis: University of Minnesota Press.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

*Heinicke, C. M., Diskin, S. D., Ramsey-Klee, D. M., & Oates, D. S. (1986). Pre- and postbirth antecedents of 2-year-old attention, capacity for relationships, and verbal expressiveness. *Developmental Psychology, 22,* 777–787.

Heise, D. R. (1969). Separating reliability and stability in test–retest correlation. *American Sociological Review, 34,* 93–101.

Hellervik, L. W., Hazucha, J. F., & Schneider, R. J. (1992). Behavior change: Models, methods, and a review of evidence. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial/organizational psychology* (Vol. 3, pp. 823–895). Palo Alto, CA: Davies-Black.

*Helson, R., & Moane, G. (1987). Personality change in women from college to midlife. *Journal of Personality and Social Psychology, 53,* 176–186.

*Helson, R., Roberts, B., & Agronick, G. (1995). Enduringness and change in creative personality and the prediction of occupational creativity. *Journal of Personality and Social Psychology, 69,* 1173–1183.

Helson, R., & Stewart, A. J. (1994). Personality change in adulthood. In T. F. Heatherton & J. L. Weinberger, *Can personality change?* (pp. 201–226). Washington, DC: American Psychological Association.

Helson, R., Stewart, A. J., & Ostrove, J. (1995). Identity in three cohorts of midlife women. *Journal of Personality and Social Psychology, 69,* 544–557.

*Helson, R., & Wink, P. (1992). Personality change in women from the early 40s to the early 50s. *Psychology and Aging, 7,* 46–55.

Higgins, E. T., & Eccles-Parsons, J. (1983). Social cognition and the social life of the child: Stages as subcultures. In E. T. Higgins, D. N. Ruble, & W. W. Hartup (Eds.), *Social cognition and social development: A sociocultural perspective* (pp. 15–62). New York: Cambridge University Press.

Hogan, R. T., & Roberts, B. W. (in press). A socioanalytic perspective on person/environment interaction. In W. B. Walsh, K. H. Craik, & R. H. Price (Eds.), *New directions in person–environment psychology.* Mahway, NJ: Erlbaum.

*Holmlund, U. (1991). Change and stability of needs from middle adolescence to young adulthood in Swedish females. *European Journal of Personality, 5,* 379–385.

*Holmlund, U. (1992). Change and stability of masculinity–femininity from adolescence to adulthood in a sample of Swedish women. *European Journal of Personality, 6,* 237–244.

Hough, L. M. (1992). The "Big Five" personality variables—Construct confusion: Description versus prediction. *Human Performance, 5,* 139–155.

*Huesmann, L. R., Eron, L. D., Lefkowitz, M. M., & Walder, L. O. (1984). Stability of aggression over time and generations. *Developmental Psychology, 20,* 1120–1134.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings.* Newbury Park, CA: Sage.

*Hunter, S. M., Johnson, C. C., Vizelberg, I. A., Webber, L. S., & Berenson, G. S. (1991). Tracking of Type A behavior in children and young adults: The Bogalusa Heart Study. *Journal of Social Behavior and Personality, 6,* 71–84.

Ickes, W., Snyder, M., & Garcia, S. (1997). Personality influences on the choice of situations. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 166–198). San Diego, CA: Academic Press.

James, W. (1950). *The principles of psychology.* New York: Dover. (Original work published in 1890)

*Jessor, R. (1983). The stability of change: Psychosocial development from adolescence to young adulthood. In D. Magnusson & V. L. Allen (Eds.), *Human development: An interactional approach* (pp. 321–341). San Diego, CA: Academic Press.

*John, O. P., Cheek, J. M., & Klohnen, E. C. (1996). On the nature of self-monitoring: Construct explication with Q-sort ratings. *Journal of Personality and Social Psychology, 71,* 763–776.

Jones, C. J., & Meredith, W. (1996). Patterns of personality change across the life span. *Psychology and Aging, 11,* 57–65.

*Kagan, J. (1960). The long term stability of selected Rorschach responses. *Journal of Consulting Psychology, 1,* 67–73.

*Kagan, J., & Moss, H. A. (1962). *Birth to maturity: A study in psychological development.* New York: Wiley.

*Kelly, E. L. (1955). Consistency of adult personality. *American Psychologist, 10,* 659–681.

*Keltikangas-Jarvinen, L. (1989). Stability of Type A behavior during adolescence, young adulthood, and adulthood. *Journal of Behavioral Medicine, 12,* 387–396.

Klohnen, E. C. (1996). Conceptual analysis and measurement of the construct of ego-resiliency. *Journal of Personality and Social Psychology, 70,* 1067–1079.

*Kochanska, G., Murray, K., & Coy, K. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68,* 263–277.

*Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandergeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67,* 490–507.

Kogan, N. (1990). Personality and aging. In J. E. Birren & S. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 330–346). San Diego, CA: Academic Press.

*Kohn, M., & Rosman, B. L. (1972). A social competence scale and symptom checklist for the preschool child: Factor dimensions, their cross-instrument generality, and longitudinal persistence. *Developmental Psychology, 6,* 430–444.

*Kohn, M., & Rosman, B. L. (1973). Cross-situational and longitudinal stability of social-emotional functioning in young children. *Child Development, 44,* 721–727.

Kohn, M. L. (1980). Job complexity and adult personality. In N. J. Smelser & E. H. Erikson (Eds.), *Themes of work and love in adulthood* (pp. 193–212). Cambridge, MA: Harvard University Press.

Kohn, M. L., & Schooler, C. (1978). The reciprocal effects of the substantive complexity of work and intellectual flexibility: A longitudinal assessment. *American Journal of Sociology, 84,* 24–52.

*Korn, S. J. (1984). Continuities and discontinuities in difficult/easy temperament: Infancy to young adulthood. *Merrill-Palmer Quarterly, 30,* 189–199.

Kremen, A. (1999). *Predicting personality in young adulthood from temperament in preschool.* Unpublished manuscript.

*Leon, G. R., Gillum, B., Gillum, R., & Gouze, M. (1979). Personality stability and change over a 30 year period: Middle age to old age. *Journal of Consulting and Clinical Psychology, 47,* 517–524.

*Lerner, J. V., Hertzog, C., Hooker, K. A., Hassibi, M., & Thomas, A. (1988). A longitudinal study of negative emotional states and adjustment from early childhood through adolescence. *Child Development, 59,* 356–366.

Levinson, D. J. (1978). *The seasons of a man's life.* New York: Ballantine.

Linehan, M. M., & Kehrer, C. A. (1993). Borderline personality disorder. In D. H. Barlow (Ed.), *Clinical handbook of psychological disorders* (2nd ed., pp. 396–441). New York: Guilford Press.

Littwin, S. (1986). *The postponed generation: Why America's grown-up kids are growing up later.* New York: Morrow.

*Loehlin, J. C., Horn, J. M., & Willerman, L. (1990). Heredity, environment, and personality change: Evidence from the Texas Adoption Project. *Journal of Personality, 58,* 221–243.

Loevinger, J. (1966). The meaning and measurement of ego development. *American Psychologist, 21,* 195–206.

*Loevinger, J., Cohn, L. D., Bonneville, L. P., Redmore, C. D., Streich, D. D., & Sargent, M. (1985). Ego development in college. *Journal of Personality and Social Psychology, 48,* 947–962.

*Lovibond, P. F. (1998). Long-term stability of depression, anxiety, and stress syndromes. *Journal of Abnormal Psychology, 107,* 520–526.

*Magnusson, D., & Backteman, G. (1978). Longitudinal stability of person characteristics: Intelligence and creativity. *Applied Psychological Measurement, 2,* 481–490.

*Martin, J., & Redmore, C. (1978). A longitudinal study of ego development. *Developmental Psychology, 14,* 189–190.

Martin, R. P., Wisenbaker, J., & Huttunen, M. (1994). Review of factor analytic studies of temperament measures based on the Thomas–Chess structural model: Implications for the Big Five. In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 157–172). Hillsdale, NJ: Erlbaum.

*Masten, A. S., Morison, P., & Pellegrini, D. S. (1985). A revised class play method of peer assessment. *Developmental Psychology, 21,* 523–533.

*Matheny, A. P., Jr. (1983). A longitudinal twin study of stability of components from Bayley's Infant Behavior Record. *Child Development, 54,* 356–360.

*Matheny, A. P., Jr. (1989). Children's behavioral inhibition over age and across situations: Genetic similarity for a trait during change. *Journal of Personality, 57,* 215–235.

McCrae, R. R., & Costa, P. T. (1990). *Personality in adulthood.* New York: Guilford Press.

McCrae, R. R., & Costa, P. T. (1994). The stability of personality: Observation and evaluations. *Current Directions in Psychological Science, 3,* 173–175.

McDevitt, S. C. (1986). Continuity and discontinuity of temperament in infancy and early childhood: A psychometric perspective. In R. Plomin & J. Dunn (Eds.), *The study of temperament: Changes, continuities and challenges* (pp. 27–38). Hillsdale, NJ: Erlbaum.

*McDevitt, S. C., & Carey, W. B. (1981). Stability of ratings vs. perceptions of temperament from early infancy to 1–3 years. *American Journal of Orthopsychiatry, 51,* 342–345.

*McGue, M., Bacon, S., & Lykken, D. T. (1993). Personality stability and change in early adulthood: A behavioral genetic analysis. *Developmental Psychology, 29,* 96–109.

McNally, S., Eisenberg, S., & Harris, J. (1991). Consistency and change in maternal child-rearing practices and values: A longitudinal study. *Child Development, 62,* 190–198.

*McNeil, T. F., & Persson-Blennow, I. (1988). Stability of temperament characteristics in childhood. *American Journal of Orthopsychiatry, 58,* 622–626.

*Melamed, A. R., Silverman, M. S., & Lewis, G. J. (1974). Three-year follow-up of women religious on the 16 Personality Factor Questionnaire. *Review of Religious Research, 15,* 64–70.

*Meyer, J. M., Heath, A. C., Eaves, L. J., Mosteller, M., & Schieken, R. M. (1988). The predictive power of Cattell's personality questionnaires: An eighteen-month prospective study. *Personality and Individual Differences, 9,* 203–212.

Mischel, W. (1968). *Personality and assessment.* New York: Wiley.

Mischel, W. (1992). Personality dispositions revisited and revised: A view after three decades. In L. Pervin (Ed.), *Handbook of personality theory and research* (pp. 111–134). New York: Guilford Press.

Modell, J. (1989). *Into one's own: From youth to adulthood in the United States 1920–1975.* Berkeley, CA: University of California Press.

*Mortimer, J. T., Finch, M. D., & Kumka, D. (1982). Persistence and change in development: The multidimensional self-concept. In P. B. Baltes & O. G. Brim (Eds.), *Life-span development and behavior* (Vol. 4, pp. 263–313). New York: Academic Press.

Moss, H. A., & Susman, E. J. (1980). Longitudinal study of personality development. In O. G. Brim, Jr., & J. Kagan (Eds.), *Constancy and change in human development* (pp. 530–595). Cambridge, MA: Harvard University Press.

*Muntaner, C., Garcia-Sevilla, L., Fernandez, A., & Torrubia, R. (1988).

Personality dimensions, schizotypal and borderline personality traits and psychosis proneness. *Personality and Individual Differences, 9,* 257–268.

*Mussen, P., Eichorn, D. H., Honzik, M. P., Bieber, S. L., & Meredith, W. M. (1980). Continuity and change in women's characteristics over four decades. *International Journal of Behavioral Development, 3,* 333–347.

Myers, I. B., & McCaulley, M. H. (1993). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator.* Palo Alto, CA: Consulting Psychologists Press.

Neprash, J. A. (1936). The reliability of questions in the Thurstone Personality Schedule. *Journal of Social Psychology, 7,* 239–244.

Nesselroade, J. R. (1991). Interindividual differences in intraindividual change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change* (pp. 92–105). Washington, DC: American Psychological Association.

Neugarten, B. L. (1968). The awareness of middle age. In B. L. Neugarten (Ed.), *Middle age and aging* (pp. 93–98). Chicago: University of Chicago Press.

*Nichols, R. C. (1967). Personality change and the college. *American Education Research, 4,* 173–190.

*Nolen-Hoeksema, S., Girgus, J. S., & Seligman, M. E. (1986). Learned helplessness in children: A longitudinal study of depression, achievement, and explanatory style. *Journal of Personality and Social Psychology, 51,* 435–442.

*O'Donnell, J. P., Leicht, D. J., Phillips, F. L., Marnett, J. P., & Horn, W. F. (1988). Stability of children's behavior problems: A 3½-year longitudinal study. *Journal of Applied Developmental Psychology, 9,* 233–241.

*Ogawa, J. R., Sroufe, L. A., Weinfield, N. S., Carlson, E. A., & Egeland, B. (1997). Development and the fragmented self: Longitudinal study of dissociative symptomatology in a nonclinical sample. *Development and Psychopathology, 9,* 855–879.

*Olson, S. L. (1989). Assessment of impulsivity in preschoolers: Cross-measure convergences, longitudinal stability, and relevance to social competence. *Journal of Clinical Child Psychology, 2,* 176–183.

*Olweus, D. (1977). Aggression and peer acceptance in adolescent boys: Two short-term longitudinal studies of ratings. *Child Development, 48,* 1301–1313.

Olweus, D. (1979). Stability of aggressive reaction patterns in males: A review. *Psychological Bulletin, 86,* 852–875.

*Ormel, J., & Schaufeli, W. B. (1991). Stability and change in psychological distress and their relationship with self-esteem and locus of control: A dynamic equilibrium model. *Journal of Personality and Social Psychology, 60,* 288–299.

Ozer, D. (1986). *Consistency in personality: A methodological framework.* New York: Springer-Verlag.

Pals, J. L. (1999). Identity consolidation in early adulthood: Relations with ego-resiliency, the context of marriage, and personality change. *Journal of Personality, 67,* 295–329.

*Pederson, W. (1991). Mental health, sensation seeking and drug use patterns: A longitudinal study. *British Journal of Addiction, 86,* 195–204.

*Pedlow, R., Sanson, A., Prior, M., & Oberklaid, F. (1993). Stability of maternally reported temperament from infancy to 8 years. *Developmental Psychology, 29,* 998–1007.

Pervin, L. A. (1994). A critical analysis of current trait theory. *Psychological Inquiry, 5,* 103–113.

*Piccione, C., Hilgard, E. R., & Zimbardo, P. G. (1989). On the degree of stability of measured hypnotizability over a 25-year period. *Journal of Personality and Social Psychology, 56,* 289–295.

*Plant, W. T. (1965a). Longitudinal changes in intolerance and authoritarianism for subjects differing in amount of college education over four years. *Genetic Psychology Monographs, 72,* 247–287.

*Plant, W. T. (1965b). Personality changes associated with college attendance. *Human Development, 8,* 142–151.

*Plant, W. T., & Telford, C. W. (1966). Changes in personality for groups completing different amounts of college over two years. *Genetic Psychology Monographs, 74,* 3–36.

*Popham, S. M., & Holden, R. R. (1991). Psychometric properties of MMPI factor scales. *Personality and Individual Differences, 12,* 513–517.

*Redmore, C. D. (1983). Ego development in the college years: Two longitudinal studies. *Journal of Youth and Adolescence, 12,* 301–306.

*Redmore, C. D., & Loevinger, J. (1979). Ego development in adolescence: Longitudinal studies. *Journal of Youth and Adolescence, 8,* 1–20.

*Reznick, J. S., Gibbons, J. L., Johnson, M. O., & McDonough, P. M. (1989). Behavioral inhibition in a normative sample. In J. S. Reznick (Ed.), *Perspectives on behavioral inhibition* (pp. 25–49). Chicago: University of Chicago Press.

Roberts, B. W. (1997). Plaster or plasticity: Are work experiences associated with personality change in women? *Journal of Personality, 65,* 205–232.

*Roberts, B. W., & Chapman, C. (in press). Change in dispositional well-being and its relation to role quality: A 30-year longitudinal study. *Journal of Research in Personality.*

Roberts, B. W., & Donahue, E. M. (1994). One personality, multiple selves: Integrating personality and social roles. *Journal of Personality, 62,* 201–218.

Roberts, B. W., & Helson, R. (1997). Changes in culture, changes in personality: The influence of individualism in a longitudinal study of women. *Journal of Personality and Social Psychology, 72,* 641–651.

Roberts, G. C., Block, J. H., & Block, J. (1984). Continuity and change in parents' child rearing practices. *Child Development, 55,* 586–597.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* Thousand Oaks, CA: Sage.

Ross, L., & Nisbett, R. E. (1991). *The person and the situation.* New York: McGraw-Hill.

Rothbart, M. K., & Bates, J. E. (1998). Temperament. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 311–388). New York: Wiley.

*Rubin, K. H., Hymel, S., & Mills, R. S. L. (1989). Sociability and social withdrawal in childhood: Stability and outcomes. *Journal of Personality, 57,* 237–255.

*Ruff, H. A., Lawson, K. R., Parrinello, R., & Weissberg, R. (1990). Long-term stability of individual differences in sustained attention in the early years. *Child Development, 61,* 60–75.

Sameroff, A. J. (1995). General systems theories and developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (pp. 659–695). New York: Wiley.

*Sanson, A., Pedlow, R., Cann, W., Prior, M., & Oberklaid, F. (1996). Shyness ratings: Stability and correlates in early childhood. *International Journal of Behavioral Development, 19,* 705–724.

Sapir, E. (1934). Personality. *Encyclopedia of the Social Sciences, 12,* 85–88.

*Saudino, K. J., & Eaton, W. O. (1995). Continuity and change in objectively assessed temperament: A longitudinal twin study of activity level. *British Journal of Developmental Psychology, 13,* 81–95.

*Schofield, W. (1953). A study of medical students with the MMPI: II. Group and individual changes after two years. *Journal of Psychology, 36,* 137–141.

Schuerger, J. M., Tait, E., & Tavernelli, M. (1982). Temporal stability of personality by questionnaire. *Journal of Personality and Social Psychology, 43,* 176–182.

Schuerger, J. M., Zarrella, K. L., & Hotz, A. S. (1989). Factors that influence the temporal stability of personality by questionnaire. *Journal of Personality and Social Psychology, 56,* 777–783.

Shiner, R. L. (1998). How shall we speak of children's personality in

middle childhood? A preliminary taxonomy. *Psychological Bulletin, 124,* 308–332.

Simonton, D. K. (1977). Eminence, creativity, and geographical marginality: A recursive structural equation model. *Journal of Personality and Social Psychology, 35,* 805–816.

*Skolnick, A. (1966). Stability and interrelations of thematic test imagery over 20 years. *Child Development, 37,* 389–396.

*Soldz, S., & Vaillant, G. E. (1999). The Big Five personality traits and the life course: A 45-year longitudinal study. *Journal of Research in Personality, 33,* 208–232.

Sroufe, L. A. (1979). Socioemotional development. In J. Osofsky (Ed.), *Handbook of infant development* (pp. 462–516). New York: Wiley.

*Stacy, A. W., Newcomb, M. D., & Bentler, P. M. (1991). Social psychological influences on sensation seeking from adolescence to adulthood. *Personality and Social Psychology Bulletin, 17,* 701–708.

*Stein, J. A., Newcomb, M. D., & Bentler, P. M. (1986). Stability and change in personality: A longitudinal study from early adolescence to young adulthood. *Journal of Research in Personality, 20,* 276–291.

*Steinberg, L. (1986). Stability (and instability) of Type A behavior from childhood to young adulthood. *Developmental Psychology, 22,* 393–402.

*Stevens, D. P., & Truss, C. V. (1985). Stability and change in adult personality over 12 and 20 years. *Developmental Psychology, 21,* 568–584.

Stewart, A. J., & Ostrove, J. M. (1998). Women's personality in middle age: Gender, history, and midcourse corrections. *American Psychologist, 53,* 1185–1194.

*Stewart, L. H. (1964). Change in personality test scores during college. *Journal of Counseling Psychology, 11,* 211–220.

*Stricker, L. J., & Ross, J. (1964). An assessment of some structural properties of the Jungian personality typology. *Journal of Abnormal and Social Psychology, 68,* 62–71.

Super, C. M., & Harkness, S. (1994). Temperament and the developmental niche. In W. B. Carey & S. C. McDevitt (Eds.), *Prevention and early intervention: Individual differences as risk factors for the mental health of children* (pp. 115–125). New York: Brunner/Mazel.

Tate, R. L., & Hokanson, J. E. (1993). Analyzing individual status and change with hierarchical linear models: Illustration with depression in college students. *Journal of Personality, 61,* 182–206.

Tellegen, A. (1988). The analysis of consistency in personality assessment. *Journal of Personality, 56,* 621–663.

Thomas, A., & Chess, S. (1977). *Temperament and development.* New York: Brunner/Mazel.

*Thomas, A., & Chess, S. (1986). The New York Longitudinal Study: From infancy to early adult life. In R. Plomin & J. Dunn (Eds.), *The study of temperament: Changes, continuities, and challenges* (pp. 39–52). Hillsdale, NJ: Erlbaum.

*Tomlinson, P. S., Harbaugh, B. L., & Anderson, K. H. (1996). Children's temperament at 3 months and 6 years old: Stability, reliability, and measurement issues. *Issues in Comprehensive Pediatric Nursing, 19,* 33–47.

*Torgersen, A. M. (1988). Genetic and environmental influences on temperament development: Longitudinal study of twins from infancy to adolescence. In S. Doxiadis (Ed.), *Early influences shaping the individual* (pp. 269–281). New York: Plenum Press.

*Troy, M. F. (1988). Antecedents, correlates, and continuity of ego-control and ego-resiliency in a high-risk sample of preschool children (Doctoral dissertation, University of Minnesota, 1988). *Dissertation Abstracts International, 49* (09), 4028.

*Tubman, J. G., Lerner, R. M., Lerner, J. V., & Von Eye, A. (1992). Temperament and adjustment in young adulthood: A 15-year longitudinal analysis. *American Journal of Orthopsychiatry, 62,* 565–574.

*Tubman, J. G., & Windle, M. (1995). Continuity of difficult temperament in adolescence: Relations with depression, life events, family support, and substance use across a one-year period. *Journal of Youth and Adolescence, 24,* 133–153.

*Tuddenham, R. D. (1959). The constancy of personality ratings over two decades. *Genetic Psychology Monographs, 60,* 3–29.

Tyler, T. R., & Schuller, R. A. (1991). Aging and attitude change. *Journal of Personality and Social Psychology, 61,* 689–697.

*Usala, P. D., & Hertzog, C. (1991). Evidence of differential stability of state and trait anxiety in adults. *Journal of Personality and Social Psychology, 60,* 471–479.

Vaillant, G. E. (1977). *Adaptation to life.* Boston: Little, Brown.

Vandewater, E. A., Ostrove, J. M., & Stewart, A. J. (1997). Predicting women's well-being in midlife: The importance of personality development and social role involvements. *Journal of Personality and Social Psychology, 72,* 1147–1160.

*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.

Visser, P. S., & Krosnick, J. A. (1998). Development of attitude strength over the life cycle: Surge and decline. *Journal of Personality and Social Psychology, 75,* 1389–1410.

*Von Dras, D. D., & Siegler, I. C. (1997). Stability in extraversion and aspects of social support at midlife. *Journal of Personality and Social Psychology, 72,* 233–241.

Wachs, T. D. (1994). Fit, context, and the transition between temperament and personality. In C. F. Halverson, Jr., G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 209–220). Hillsdale, NJ: Erlbaum.

West, S. G., & Graziano, W. G. (1989). Long-term stability and change in personality: An introduction. *Journal of Personality, 57,* 175–193.

*Westenberg, P. M., & Gjerde, P. F. (1999). Ego development during the transition from adolescence to young adulthood: A 9-year longitudinal study. *Journal of Research in Personality, 33,* 233–252.

*Wheeler, D. S., & Schwartz, J. C. (1989). Millon Clinical Multiaxial Inventory (MCMI) scores with a collegiate sample: Long-term stability and self–other agreement. *Journal of Psychopathology and Behavioral Assessment, 11,* 339–352.

*Whitbourne, S. K., Zuschlag, M. K., Elliot, L. B., & Waterman, A. S. (1992). Psychosocial development in adulthood: A 22-year sequential study. *Journal of Personality and Social Psychology, 63,* 260–271.

*Wiggins, J. S., & Winder, C. L. (1961). The Peer Nomination Inventory: An empirically derived sociometric measure of adjustment in preadolescent boys. *Psychological Reports, 9,* 643–677.

*Wilhelm, K., & Parker, G. (1990). Reliability of the Parental Bonding Instrument and Intimate Bond Measure scales. *Australian and New Zealand Journal of Psychiatry, 24,* 199–202.

Winter, D. G. (1988). The power motive in women—and men. *Journal of Personality and Social Psychology, 54,* 510–519.

*Woodall, K. L., & Matthews, K. A. (1993). Changes in and stability of hostile characteristics: Results from a 4-year longitudinal study of children. *Journal of Personality and Social Psychology, 64,* 491–499.

*Woodruff, D. S. (1983). The role of memory in personality continuity: A 25 year follow-up. *Experimental Aging Research, 9,* 31–34.

*Yonge, G. D., & Regan, M. C. (1975). A longitudinal study of personality and choice of major. *Journal of Vocational Behavior, 7,* 41–65.

Young, C. H., Savola, K. L., & Phelps, E. (1991). *Inventory of longitudinal studies in the social sciences.* London: Sage.

Received March 20, 1998
Revision received June 17, 1999
Accepted July 13, 1999 ■