Journal of Personality and Social Psychology
2003, Vol. 84, No. 1, 205–220

0022-3514/03/$12.00 DOI: 10.1037/0022-3514.84.1.205

Stability of Self-Esteem Across the Life Span

Kali H. Trzesniewski, M. Brent Donnellan, and Richard W. Robins
University of California, Davis

Two studies examined the rank-order stability of self-esteem from age 6 to 83: Study 1 was a meta-analysis of 50 published articles (N = 29,839) and Study 2 analyzed data from 4 large national studies (N = 74,381). Self-esteem showed substantial continuity over time (disattenuated correlations ranged from the .50s to .70s), comparable to the stability found for personality traits. Both studies provided evidence for a robust developmental trend: Self-esteem stability was low during childhood, increased throughout adolescence and young adulthood, and declined during midlife and old age. This trend could not be explained by age differences in the reliability of self-esteem measures, and generally replicated across gender, ethnicity, self-esteem scale, nationality (U.S. vs. non-U.S.), and year of publication.

Global self-esteem is one of the most frequently studied variables in psychology. Research has linked high self-esteem to many positive outcomes, including occupational success (e.g., Bachman & O'Malley, 1977; Elliott, 1996; Judge & Bono, 2001; Mortimer, Finch, & Kumba, 1982), healthy social relationships (e.g., Hendrick, Hendrick, & Adler, 1988; Leary, Tambor, Terdal, & Downs, 1995; Murray, Holmes, & Griffin, 2000; Neyer & Asendorpf, 2001), subjective well-being (e.g., Campbell, 1981; DeNeve & Cooper, 1998; Diener & Diener, 1995; Fordyce, 1988; Wilson, 1967), positive perceptions by peers (e.g., Paulhus, 1998; Robins, Hendin, & Trzesniewski, 2001), academic achievement (e.g., Hansford & Hattie, 1982; Liu, Kaplan, & Risser, 1992; Rubin, 1978), persistence in the face of failure (e.g., Cruz-Perez, 1973; McFarlin, Baumeister, & Blascovich, 1984; Shrauger & Sorman, 1977), and improved coping and self-regulation skills (e.g., Greenberg et al., 1999; Heatherton, Herman, & Polivy, 1991; Musser & Browne, 1991). Conversely, low self-esteem has been linked to a number of problematic outcomes, including depressive symptoms (e.g., Block, Gjerde, & Block, 1991; Reinherz, Giaconia, Pakiz, & Silverman, 1993; J. E. Roberts, Gotlib, & Kassel, 1996; Whisman & Kwon, 1993), health problems (e.g., Antonucci & Jackson, 1983; O'Connor & Vallerand, 1998; Vingilis, Wade, & Adlaf, 1998), and antisocial behavior (e.g., Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2002; Owens, 1994; Rosenberg, Schooler, & Schoenbach, 1989).[1]

Researchers who study the correlates and consequences of self-esteem generally assume it is a stable trait that predicts future behavior (e.g., Harter, 1998; Rosenberg, 1965). However, this assumption has been questioned by researchers who emphasize the state-like nature of self-esteem. In Conley's (1984) influential review of personality stability, he argued that self-esteem "cannot be considered a stable individual-differences construct" because it lacks long-term stability (p. 21). He further argued that self-esteem is highly influenced by the individual's environment, which undermines its stability and limits its ability to predict future behavior.

Similarly, many self-esteem theorists believe that the construct is better conceptualized in terms of state-like processes. Leary and Baumeister (2000) characterized self-esteem as a barometer of transient beliefs about one's worth relative to others. From this perspective, self-esteem levels are highly reactive to social evaluation and are thus continually changing in response to external feedback. Kernis and his colleagues (e.g., Kernis, Cornell, Sun, Berry, & Harlow, 1993; Kernis, Grannemann, & Barclay, 1992) emphasized the importance of daily fluctuations in self-esteem and argued that the degree to which a person's self-esteem fluctuates on a day-to-day basis predicts future behavior independently of a person's average level of self-esteem over time. Experimental research has linked a number of self-evaluative processes to short-term changes in self-esteem (e.g., Brown & Mankowski, 1993; Gergen, 1981; Heatherton & Polivy, 1991; Jones, Rhodewalt, Berglas, & Skelton, 1981; Markus & Kunda, 1986; Tesser, 1988), leading some researchers to conclude that people's self-esteem levels can be moved around quite easily. Reflecting these various perspectives, Heatherton and Polivy (1991) developed a state self-esteem scale for use in research on short-term changes in self-esteem, and Rosenberg (1965) created a scale that assesses the degree to which a person's self-esteem remains stable over time.

A key question in the state-trait debate concerns the actual stability of self-esteem over time. To address this question, the

Kali H. Trzesniewski and Richard W. Robins, Department of Psychology, University of California, Davis; M. Brent Donnellan, Department of Human and Community Development, University of California, Davis.

This research was supported by National Institute of Mental Health Grant MH61829 and by a grant from the University of California, Davis. We thank R. Chris Fraley, Erika E. Gaylor, Simona Ghetti, Katie L. Isaacson, Brent W. Roberts, Jessica L. Tracy, and Phillip R. Shaver for their helpful comments on an earlier version of this article. We thank Keith F. Widaman for his statistical advice and Christine Silvani for help compiling articles for Study 1.

Correspondence concerning this article should be addressed to Kali H. Trzesniewski, Department of Psychology, University of California, Davis, California 95616-8686. E-mail: kftrzesniewski@ucdavis.edu

[1] It is important to note that some researchers have questioned both the benefits of high self-esteem and the costs of low self-esteem (e.g., Baumeister, Smart & Boden, 1996).

present research analyzes longitudinal data to determine the degree to which self-esteem shows trait-like stability over long periods of time. A better understanding of the state-trait distinction has important implications for theories of self-esteem. For example, should self-esteem be conceptualized as a trait and, if so, is it a highly stable trait like intelligence or a moderately stable trait like personality? That is, where exactly does self-esteem fall on the hierarchy of consistency (e.g., Conley, 1984; Kelly, 1955)? If self-esteem is found to be less stable than intelligence and personality, should it be conceptualized as a state such as mood or is it best conceptualized as something that has both state and trait components such as depression (Dumenci & Windle, 1996; Kenny & Zautra, 2001)?

Another theoretical issue concerns whether there are periods in the life span when self-esteem exhibits greater or lesser stability. For example, is early adolescence a time of storm and stress accompanied by a dramatic transformation of the self or a period of relative continuity in self-esteem? Is self-esteem set like plaster after the age of 30 or does it get more stable with age without ever completely crystallizing, much like personality traits (e.g., B. W. Roberts & DelVecchio, 2000).

In addition to these theoretical issues, a better understanding of the stability of self-esteem would provide information about the potential real-world consequences of self-esteem, as well as the likelihood that situational and environmental factors will produce changes in self-esteem. If self-esteem exhibits very low stability over time, this would challenge the notion that it is important for long-term psychosocial outcomes. Knowledge about when self-esteem is most and least stable has implications for the timing of self-esteem interventions. Developmental periods during which stability is relatively low may be easier targets of intervention programs. These periods of relative upheaval in the self-concept may be the best times to promote positive self-evaluations because self-esteem levels may be particularly malleable.

Despite the importance of these issues, we know surprisingly little about the stability of self-esteem across the life span, compared with what we know about the stability of other personality characteristics (B. W. Roberts & DelVecchio, 2000). Few studies have focused explicitly on understanding the stability of self-esteem over long periods of time. Most longitudinal studies of self-esteem report information on stability (i.e., test–retest correlations) as background information for other research questions.

The present research examined three questions about the stability of self-esteem: (a) On average, how stable is self-esteem across the life span, and is the level of stability as high as that found for personality traits?; (b) Does the stability of self-esteem vary across developmental periods and, if so, at what age does stability peak?; and (c) Does the level of stability vary across gender, ethnicity, self-esteem measure, nationality (U.S. vs. non-U.S.), and the year the study was conducted? We conducted two studies to address these questions. In Study 1, we report findings from a meta-analysis of existing longitudinal studies of self-esteem stability. In Study 2, we analyzed longitudinal data from four large national studies.

Defining Stability

The present research focuses on the rank-order stability of self-esteem. Rank-order stability is typically assessed using test–retest correlations (i.e., the correlation between self-esteem scores across two time points). Test–retest correlations reflect the degree to which the relative ordering of individuals is maintained over time. A high test–retest correlation indicates that an individual who is high (or low) in self-esteem relative to others at one point in time remains high (or low) in self-esteem relative to others at another point time.

Rank-order stability is conceptually and statistically distinct from mean-level change (e.g., Caspi & Roberts, 1999; Robins, Fraley, Roberts, & Trzesniewski, 2001). For example, the rank ordering of individuals in a sample could change substantially over time without producing any aggregate increases or decreases in mean-levels of self-esteem (e.g., if the number of people who decreased offsets the number of people who increased). Similarly, individuals in a sample could increase substantially in self-esteem across time but the rank ordering of individuals would be maintained if everyone increased by the same amount.

Thus, a high test–retest correlation indicates that (a) individuals do not change much over time or (b) individuals are changing over time, but in more or less the same way (i.e., everyone is increasing or decreasing to the same extent). The latter situation can occur when a normative developmental event such as puberty impacts all individuals in the same way (e.g., if puberty causes everyone to decline in self-esteem by the same amount). A low test–retest correlation indicates that (a) individuals are changing over time and (b) there are individual differences in the direction of change; that is, some individuals are increasing in self-esteem whereas others are decreasing. A low test–retest correlation can occur when nonnormative developmental events impact self-esteem (e.g., if some individuals experience parental divorce and decline in self-esteem whereas others do not experience parental divorce and maintain the same self-esteem level). A low test–retest correlation can also occur when the factors that influence self-esteem are normative but individuals have unique reactions to these events (e.g., if puberty causes some individuals to increase in self-esteem but causes others to decrease in self-esteem). Finally, a low test–retest correlation could simply reflect measurement error; less reliable self-esteem scales will show lower test–retest correlations.

Previous Research on Self-Esteem Stability

Only a few studies have examined the rank-order stability of self-esteem during childhood. Marsh, Craven, and Debus (1998) reported 1-year test–retest correlations around .30 in a sample of young children (ages 5 to 7). Somewhat higher stabilities have been found for older children: Granleese and Joseph (1994) reported a 3-year test–retest correlation of .61 for a group of children followed from age 8 to 11; Rubin (1978) reported a 3-year test–retest correlation of .43 from age 9 to 12; and Nottleman (1987) reported a 1-year test–retest correlation of .51 from age 11 to 12.

In research on adolescents and young adults, test–retest correlations are typically around .50 over 1- to 6-year time intervals (e.g., Alsaker & Olweus, 1992; Bachman & O'Malley, 1977; Block & Robins, 1993; Elliot, 1996; Hirsch & DuBois, 1991; R. E. Roberts & Bengtson, 1996; Roeser & Eccles, 1998; Stein, Newcomb, & Bentler, 1986). In these studies, stability seems to increase with age and tends to be higher over shorter periods of time, reflecting a simplex pattern that is typical in longitudinal studies of trait stability.

In research on adults, test–retest correlations have ranged in magnitude even more dramatically. For example, Helson and Moane (1987) reported a test–retest correlation of .72 from age 27 to 43. However, other studies have found much lower test–retest correlations. For example, Reitzes, Mutran, and Fernandez (1996) reported a test–retest correlation of .47 from age 61 to 63, and Mussen, Eichorn, Honzik, Bieber, and Meredith (1980) found a test–retest correlation of .35 from age 30 to 70.

In summary, childhood correlations ranged from .30 to .60, adolescent and early adult correlations were generally in the .40s and .50s, and adulthood correlations ranged from .35 to .72. However, this qualitative review does not provide a precise picture of how self-esteem stability varies with age, and the inconsistent results make it difficult to draw any definite conclusions. Moreover, the large range for the correlations in both the childhood and adult samples suggests the existence of moderator variables (e.g., gender) that might account for the variation in self-esteem stability across studies. For example, Block and Robins (1993) found that adolescent girls showed significantly higher self-esteem stability (r = .55) than adolescent boys (r = .25). In the present research, we addressed these possibilities by exploring moderator variables in a meta-analysis of the existing literature (Study 1) and in several longitudinal studies based on large, heterogeneous samples (Study 2).

Study 1

Study 1 reports findings from a meta-analysis of self-esteem stability. Meta-analytic techniques provide a way to quantitatively and systematically aggregate the results from the longitudinal studies of self-esteem that have accumulated over the past several decades. This study had three general aims. Our first aim was to document the overall level of stability of self-esteem across the life span. Our second aim was to examine whether the level of stability varies as a function of the (a) age, (b) gender, (c) measure of self-esteem used in the study, (d) U.S. versus non-U.S. participants, and (e) year of assessment. Our third aim was to identify gaps and limitations in the existing research literature.

Method

Identification of Relevant Studies

We used five methods to locate relevant studies. First, we reviewed reference lists from previously published reviews of self-concept development (Demo, 1992; Harter, 1982, 1998; Wylie, 1979) and from two recent meta-analyses of gender differences in self-esteem (Kling, Hyde, Showers, & Buswell, 1999; Major, Barr, Zubek, & Babey, 1999). Second, we searched the *PsycINFO* and *ERIC* databases for articles published between 1887 (the earliest entry in the PsycINFO database) and June 2002, using the keyword "self-esteem" paired with each of the following keywords: age differences, change, consistency, continuity, development, literature review, longitudinal, meta-analysis, and stability. Third, we paired the keyword "self-esteem" with every journal title that contained a previously identified article (from the keyword search). Fourth, we searched *PsycINFO* using the names of common self-esteem scales as the keyword (e.g., Rosenberg Self-Esteem Scale, Self-Esteem Inventory, Piers–Harris Self-Concept Scale). These keyword searches resulted in 9,410 citations (7,150 were from peer-reviewed journals). Fifth, we searched for relevant articles by reviewing the reference lists of the articles identified in the *PsycINFO* and *ERIC* searches that met the inclusion criteria described in the next section.

Criteria for Inclusion of Study

We included studies if they fulfilled six criteria. First, the study had to be published in a peer-reviewed journal. Second, the study had to include a measure of global self-esteem; measures of domain-specific self-evaluations (e.g., academic self-esteem, body image) were not included.[2] Third, the study had to report test–retest correlations (i.e., the correlation between two time points) and specify the age of participants (or an age range) at each time point. Fourth, the study had to examine test–retest correlations across at least a 1-year interval. This criterion emphasized the long-term stability of self-esteem and diminished potential carry-over effects that may have inflated stability estimates. Fifth, the sample had to be drawn from a nonclinical population. Sixth, the study could not involve an intervention or manipulation. These last two criteria ensured that studies were reporting normative information about self-esteem stability. Fifty studies satisfied these criteria, resulting in a total of 168 test–retest correlations based on 29,839 participants.[3]

Study Variables

Rank-order stability. Rank-order stability was measured by test–retest correlations (i.e., the correlation of a self-esteem scale administered at two points in time).

Age. The age for each measurement point was recorded. Seventeen studies reported only an age range (e.g., 20 to 30). For these studies, we used the median or mean of the age range if this information was provided; otherwise, we used the midpoint of the reported age range as an estimate (e.g., 25 was used if the reported age range was 20 to 30).

Time interval. For each test–retest correlation, the time interval between assessments was coded.

Gender. If test–retest correlations were reported by gender, this information was recorded separately for men and women.

Self-esteem measure. The particular self-esteem scale used in each study was recorded. We created three categories for the Rosenberg (1965) Self-Esteem Scale: (a) the original 10-item scale, (b) abbreviated versions of the original scale, and (c) adapted versions of the original scale (i.e., a combination of original and new items). Four studies used more than one measure of self-esteem; therefore test–retest correlations were recorded for a total of 54 measures.

Sample size. Sample size was recorded by age for each study. If a study reported test–retest correlations by gender, then the sample size by gender was also recorded.

Reliability. Cronbach's coefficient alpha was recorded for each assessment point. Five studies reported an average alpha reliability across assessments. For these studies, the average reliability was coded for all assessment points.

Nationality. The country in which the data were collected was coded as either the United States or outside of the United States.

Year of assessment. The year the data were collected was recorded. Twenty-two studies did not report the year of data collection. For these studies, the year of data collection was estimated as 9 years prior to the date of publication (the median time lag between data collection date and publication date based on the studies that provided both).

---

[2] Global self-esteem is conceptually distinct from domain-specific self-evaluations in that the focus is on the whole self as opposed to a specific facet of the self, such as academic competence (see Rosenberg, Schooler, Schoenbach, & Rosenberg, 1995).

[3] A list of the articles that were included in the meta-analysis is available from Kali H. Trzesniewski.

Results and Discussion

Study Characteristics

For the 50 studies in the meta-analytic database, the median age of participants at initiation of the study was 14.0 ($SD = 14.0$). The average study spanned a total of 4.9 years ($SD = 6.7$; range = 1 to 40 years); the average time interval between individual assessments was 3.2 years ($SD = 4.4$). Thirty-five studies (70%) were based on two waves of data, 11 (22%) were based on three waves, 3 (6%) were based on four waves, and 1 (2%) was based on five waves. The median number of participants per study was 258 (range = 24 to 4,591). The mean alpha reliabilities were .78 for Wave 1, .78 for Wave 2, .75 for Wave 3, .80 for Wave 4, and .81 for Wave 5. These reliabilities were used to compute disattenuated test–retest correlations (i.e., correlations corrected for measurement error).[4]

The most striking finding was the dearth of studies examining self-esteem in childhood, adulthood, or old age. The vast majority of studies (66%) included participants in the preadolescent and adolescent age groups (ages 11 to 19). Only 13% of the studies examined self-esteem prior to age 11 and only 21% examined self-esteem past age 19. Even more striking, only 5% of the studies examined self-esteem after age 40.

In terms of the gender composition of the studies, 4% included men only, 6% included women only, and 90% included both men and women. However, only 24% of the studies reported test–retest correlations by gender because most of the mixed gender studies reported results only for the total sample. Thus, in about three fourths of the studies, it was impossible to determine whether the level of stability was similar for men and women.

Another noteworthy finding is the lack of consistency in the way self-esteem has been measured. No single measure was used in a majority of the studies. In fact, 30% of the studies used a self-esteem measure that was included only in that particular study. The most commonly used measure was the Rosenberg (1965) Self-Esteem Scale (20% used the original scale, 11% used an abbreviated version, and 6% used an adapted version), followed by the Coopersmith (1967) Self-Esteem Inventory (14%), the Global Self-Worth subscale of the Self-Perception Profile for Children (15%; Harter, 1985), and the Self-Description Questionnaire (6%; Marsh, Smith, & Barnes, 1983). The use of different measures reflects, in part, the fact that some scales were developed for use with particular age groups. For example, Harter's Global Self-Worth scale was developed for use with children and adolescents and was used only in studies of participants younger than 18.

Age Differences in Self-Esteem Stability

We followed the approach of Hedges and Olkin (1985) to compute estimates of self-esteem stability (see also B. W. Roberts & DelVecchio, 2000). Population estimates were created by converting test–retest correlations using Fisher's *r*-to-*z* transformation and then weighting the transformed effect sizes by the inverse of their variance.[5] Confidence intervals and tests of heterogeneity were calculated using formulas from Hedges and Olkin. Following B. W. Roberts and DelVechhio (2000), we used analysis of covariance (ANCOVA) to reestimate population estimates of self-esteem stability controlling for relevant covariates (e.g., time interval).

The overall mean test–retest correlation, averaged across all studies and time periods was .50 ($N = 168$ test–retest correlations). The mean disattenuated test–retest correlation was .64. Table 1 shows the year-by-year estimated population test–retest correlations from ages 6 to 82. We found statistically significant heterogeneity in the estimated population correlations for most ages, indicating that there was variability in test–retest correlations across the studies included in the meta-analysis. This finding suggests the presence of moderator variables.

Consistent with previous research, we found that self-esteem stability was lower over longer periods of time, as indicated by a negative relation between time interval and stability ($r = -.31$, $p < .05$). It is possible that the effect of time interval is conflated with the effect of age. To address this issue, we used multiple regression analyses to reestimate the relation between time interval and stability controlling for age, and found that time interval had an independent effect on stability ($\beta = -.35$, $p < .01$). To illustrate this effect, we used the unstandardized beta weight estimates from the regression equation including age and time interval as predictors of self-esteem stability and estimated the average stability one could expect for various lengths of time, holding age constant at 16.2 years (the mean age of the sample, weighted by sample size). The average estimated stability was .50 over a 1-year period, .46 over a 5-year period, .41 over a 10-year period, and .31 over a 20-year period.

Given the relation between time interval and stability, it is possible that comparisons across age groups are confounded by differences in the average time interval. For example, stability may appear to be lower during a particular age group simply because studies conducted on that age group tend to have longer time intervals than studies conducted on other age groups. To test this possibility, we reestimated the population test–retest correlations depicted in Table 1 using time interval as a covariate in an ANCOVA model. This analysis is equivalent to estimating the test–retest estimates as if all studies lasted for 2.9 years (the mean time interval, weighted by sample size). These new population estimates controlling for time interval are shown in the third column of Table 1. In general, these estimates were very close to the estimated population correlations without controlling for time interval.

During childhood, adulthood, and old age, most of the year-by-year test–retest correlations were based on only one study (and some of these studies had relatively small samples). As a result, the confidence intervals for these periods are large and the correlations fluctuate substantially from year to year. Following B. W. Roberts and DelVelcchio's (2000) study, we created age categories corresponding to middle childhood (ages 6 to 11), adolescence (ages 12

[4] The formula for disattenuating correlations is based on the assumptions of classical test theory. To the extent that these assumptions are violated, the resulting disattenuated correlations will be biased. In particular, the assumption that error is completely random (i.e., that there is no correlated error across time) may not be true for self-esteem measures. Therefore, the disattentuation formula is likely to produce positively biased estimates and the "true" stability of self-esteem may lie somewhere between the raw and disattenuated test–retest correlations (Marsh & Grayson, 1994).

[5] All subsequent analyses were conducted on correlations transformed using Fisher's *r*-to-*z* formula.

Table 1
Study 1: Population Estimates of Self-Esteem Stability by Age

Age at Time 1 (in years)	ρ	ρ_t	K	C. I.	N	T	Q
6	.27	.22	1	.10 to .42	127	1.00	0.01
7	.25	.20	1	.09 to .40	139	1.00	0.00
8	.40	.38	3	.30 to .50	264	2.02	2.98
9	.40	.39	3	.29 to .50	236	2.68	2.37
10	.39	.45	2	.16 to .57	70	5.50	1.85
11	.45	.41	13	.42 to .48	2,973	1.64	44.37*
12	.46	.42	35	.45 to .47	11,670	1.80	185.54*
13	.50	.46	26	.48 to .51	7,716	1.81	138.22*
14	.50	.49	23	.47 to .52	4,987	2.91	180.78*
15	.61	.59	11	.58 to .64	1,759	2.43	43.83*
16	.51	.48	6	.49 to .52	11,763	2.25	221.99*
17	.52	.50	7	.51 to .54	7,645	2.30	133.73*
18	.53	.53	10	.51 to .55	4,468	3.30	129.83*
19	.47	.60	7	.45 to .48	6,083	7.39	50.46*
20	.53	.51	1	.38 to .65	109	2.00	0.38
21	.68	.62	3	.52 to .80	68	1.00	1.77
22	.70	.68	1	.63 to .76	256	2.00	7.08*
23	.64	.72	2	.60 to .68	757	5.89	33.66*
25	.49	.52	1	.42 to .55	489	4.00	1.03
27	.72	1.00	1	.60 to .81	81	16.00	2.75
30	.35	1.00	1	.09 to .57	53	40.00	0.01
32	.64	.59	1	.50 to .75	94	1.00	1.27
34	.54	.49	1	.38 to .67	96	1.00	0.38
36	.47	.47	1	.37 to .56	273	3.00	0.43
39	.64	.74	2	.57 to .70	307	6.15	4.28
61	.47	.45	1	.41 to .52	757	2.00	1.21
69	.39	.72	1	.17 to .57	72	14.00	0.03
71	.37	.63	2	.23 to .49	170	11.81	0.12
82	.47	.45	1	.26 to .64	69	2.00	0.11

Note. ρ = estimated population correlation; ρ_t = estimated population correlation controlling for time interval; K = number of samples; C. I. = 95% confidence interval of population correlation; N = total number of participants; T = time interval (weighted by sample size); Q = heterogeneity statistic.
* $p < .05$.

to 17), and the college years (ages 18 to 21). Beyond the college years we created age categories based on decades of life.

Figure 1 shows the estimated population test–retest correlations, controlling for time interval, as a function of age group. Examination of the confidence intervals around each estimated mean revealed several noteworthy findings.[6] First, self-esteem stability was significantly lower during childhood (age 6 to 11) than at any other age except late adulthood (age 60 to 82). We found the same pattern for the disattenuated test–retest correlations, indicating that the age effects were not due to age differences in the reliability of self-esteem measures.[7] Second, self-esteem stability was significantly higher during early adulthood (age 22 to 29) than at any other age except the subsequent age period (30 to 39). Thus, self-esteem stability appears to peak in young adulthood. Again, this age effect was not due to age differences in reliability. Third, the disattenuated test–retest correlations showed that self-esteem peaks below unity ($\rho = .71$ for the 22 to 29 age group).

Visual inspection of Figure 1 suggests a curvilinear trend, such that self-esteem stability increases from childhood to adulthood and then decreases from adulthood to old age. To test this trend, we conducted a hierarchical multiple regression analysis predicting self-esteem test–retest correlations from time interval and age modeled as a linear (age), quadratic (age^2), and cubic (age^3) function. Consistent with a curvilinear trend, we found that age^2 accounted for 13% ($p < .05$) of the variance in uncorrected test–retest correlations and 4% ($p < .05$) of the variance in the disattenuated test–retest correlations; age and age^3 had weaker effects.

We also tested whether the curvilinear trend was moderated by any of the other variables included in the meta-analysis; specifically, we tested for interactions between these variables and quadratic age. For gender, nationality, and year of assessment, none of the three interaction effects was significant and none accounted for more than 1% of the variance in either the uncorrected or disattenuated test–retest correlations. To test whether the curvilinear trend held across self-esteem scales, we entered a dummy code for each scale representing the interaction between that scale and quadratic age. None of the interaction effects was significant and

[6] To determine whether two test–retest correlations were significantly different, we examined whether their 95% confidence intervals overlapped (Hedges & Olkin, 1985).

[7] Only one of the five studies from the age 30–39 group reported an alpha reliability ($\alpha = .73$). The other four studies used the original Rosenberg Self-Esteem Scale. For these studies, we estimated the alpha reliability to be .89, based on the mean alpha coefficient for all other studies of adulthood that used the original Rosenberg Self-Esteem Scale.

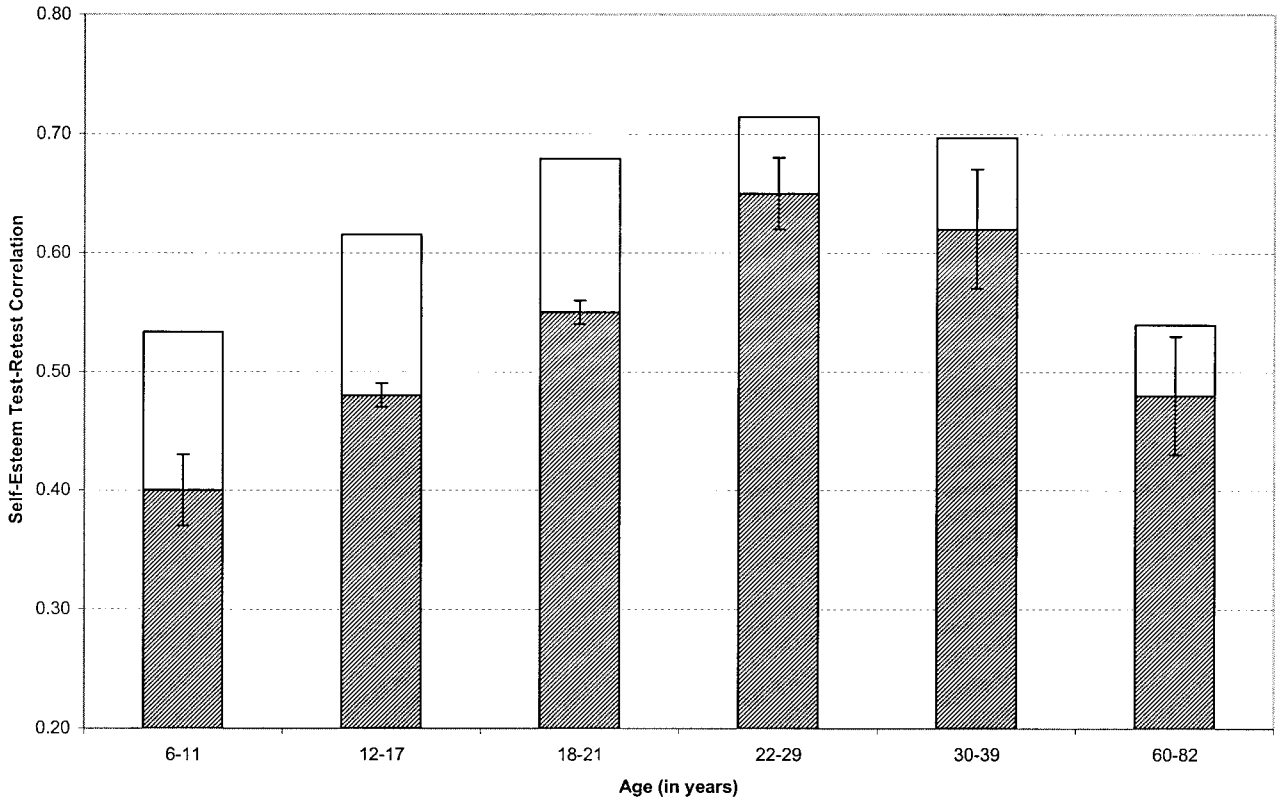

Figure 1. Study 1: Self-esteem stability as a function of age group. Values are population estimates of self-esteem stability, controlling for time interval. Shaded bars represent uncorrected test–retest correlations; nonshaded bars represent disattenuated test–retest correlations. T-bars indicate 95% confidence intervals for each age group.

none accounted for more than 2% of the variance in either the uncorrected or disattenuated test–retest correlations; when entered as a set, the scale interaction effects collectively accounted for 3.8% (*ns*) of the variance in the uncorrected correlations and 1.7% (*ns*) of the variance in the disattenuated correlations. Overall, these findings suggest that the quadratic function fit the age trends equally well for men and women, for U.S. and non-U.S. participants, and for studies conducted in different decades and using different self-esteem scales.

Gender Differences in Self-Esteem Stability

We next examined the average stability for male, female, and mixed gender samples. Most of the age categories consisted of a small number of studies, so we collapsed the data across age categories and focused on the stability estimates controlling for the effects of time interval and age (see column 4 of Table 2). Self-esteem stability did not differ by gender: The average level of stability was exactly the same ($\rho = .49$) for mixed gender samples, female samples, and male samples. The disattenuated test–retest correlations also revealed very slight gender differences, none of which were significant (see column 5 of Table 2). In summary, we found no evidence for gender differences in self-esteem stability. However, because less than one quarter of the studies reported test–retest correlations by gender, the conclusions that can be drawn from these analyses are limited.

Scale Differences in Self-Esteem Stability

Self-esteem stability did not vary substantially across self-esteem scales (see column 4 of Table 2). The only statistically significant difference was that the three versions of the Rosenberg Self-Esteem Scale ($\rho = .50$ for original and adapted; $\rho = .51$ for abbreviated) and the Miscellaneous scales category ($\rho = .52$) had significantly higher stabilities than the Coopersmith Self-Esteem Inventory ($\rho = .45$), the Harter Global Self-Worth Scale ($\rho = .43$), and the Self-Description Questionnaire ($\rho = .43$). An examination of the disattenuated test–retest correlations (see column 5 of Table 2) revealed a similar pattern, except that the Rosenberg Self-Esteem Scale (original version) no longer showed higher stability than the Coopersmith Self-Esteem Inventory, Harter Global Self-Worth Scale, and Self-Descriptive Questionnaire.

Nationality Differences in Self-Esteem Stability

To compare the average stability for U.S. and non-U.S. participants, we collapsed the data across age categories and controlled for the effects of time interval and age (see column 4 of Table 2).

Table 2
Study 1: Population Estimates of Self-Esteem Stability by Gender, Self-Esteem Measure, Nationality, and Year of Assessment

	ρ	ρ_d	ρ_{ta}	ρ_{dta}	*K*	C. I.	*N*	*T*	*Q*
Gender									
Males & females reported together	.51	.65	.49	.62	62	.50 to .52	24,546	2.26	558.35*
Males only	.50	.64	.49	.64	58	.49 to .51	26,567	3.01	559.82*
Females only	.47	.62	.49	.63	48	.46 to .50	12,438	3.77	277.25*
Self-esteem measure									
Coopersmith Self-Esteem Inventory	.45	.59	.45	.59	21	.44 to .47	8,803	2.55	138.79*
Harter Global Self-Worth Scale	.46	.61	.43	.57	14	.44 to .47	7,847	1.24	61.67*
Rosenberg Self-Esteem Scale (original)	.51	.59	.50	.57	22	.49 to .53	4,101	3.91	77.45*
Rosenberg Self-Esteem Scale (abbreviated)	.53	.68	.51	.65	25	.51 to .56	3,662	2.01	111.34*
Rosenberg Self-Esteem Scale (adapted)	.49	.61	.50	.63	12	.48 to .50	19,872	3.72	381.79*
Marsh Self-Description Questionnaire	.44	.56	.43	.54	5	.39 to .49	1,026	2.43	30.80*
Miscellaneous[a]	.53	.69	.52	.68	69	.52 to .54	18,240	2.76	526.75*
Nationality									
United States	.48	.62	.50	.64	94	.47 to .49	49,098	3.13	1,058.18*
Outside United States	.51	.63	.49	.60	74	.49 to .53	14,453	2.00	414.41*
Year of assessment									
1950s	.61	.90	.65	.96	5	.51 to .68	230	6.93	7.71
1960s	.50	.64	.51	.65	24	.49 to .51	18,545	3.57	404.36*
1970s	.41	.58	.46	.65	27	.40 to .43	8,217	6.39	105.11*
1980s	.49	.62	.47	.59	98	.48 to .50	27,390	1.74	598.24*
1990s	.57	.69	.54	.65	14	.55 to .58	9,169	1.58	243.99*

Note. ρ = estimated population correlation; ρ_d = disattenuated estimated population correlation; ρ_{ta} = estimated population correlation controlling for time interval and age; ρ_{dta} = disattenuated estimated population correlation controlling for time interval and age; *K* = number of samples; C. I. = 95% confidence interval of population correlation; *N* = total number of participants; *T* = time interval (weighted by sample size); *Q* = heterogeneity statistic.
[a] Measures included in the Miscellaneous category were used in only one study.
* $p < .05$.

There was no difference between U.S. participants ($\rho = .50$) and non-U.S. participants ($\rho = .49$). However, for the disattenuated test–retest correlations (see column 5 of Table 2), U.S. participants ($\rho = .64$) showed slightly higher levels of stability than non-U.S. participants ($\rho = .60$).

Year of Assessment Differences in Self-Esteem Stability

To examine the effect of year of assessment on self-esteem stability, we collapsed the data across age categories and controlled for the effect of time interval and age (see column 4 of Table 2). Although there was no linear relation between self-esteem stability and year of assessment, we did find some significant differences; self-esteem stability was higher for studies conducted in the 1950s ($\rho = .65$) than for studies conducted in the 1960s ($\rho = .51$) and 1990s ($\rho = .54$), which had higher stability than studies conducted in the 1970s ($\rho = .46$) and 1980s ($\rho = .47$). The disattenuated test–retest correlations revealed a similar ordering (see column 5 of Table 2), except that studies conducted in the 1970s no longer differed significantly from those conducted in the 1960s and 1990s.

In summary, the results of the meta-analysis suggest that self-esteem stability increases from childhood to early adulthood, and then decreases during middle and late adulthood. Thus, our findings show that early adulthood represents the apex of self-esteem stability whereas B. W. Roberts and DelVecchhio (2000) found that personality stability continues to increase throughout the life span (although it begins to plateau in middle adulthood). These age trends held even after controlling for age differences in time interval and scale reliability. Furthermore, the highest level of stability was well below unity, suggesting that self-esteem continues to change throughout the life span. Finally, we found that self-esteem stability did not vary substantially by gender, self-esteem scale, nationality, or year of assessment.

Study 1 has several limitations. First, the meta-analytic findings provide little information about the stability of self-esteem during adulthood or old age, because of the paucity of studies beyond adolescence. Second, many of the existing studies were based on small, homogeneous, and nonrepresentative samples. This may have led to a restriction of range in self-esteem levels, which would tend to decrease stability. Thus, although our meta-analysis points to some potentially important developmental trends, the results do not provide a high-resolution picture of the stability of self-esteem over the entire life span.

Study 2

Study 2 replicates and extends the findings of the meta-analysis by examining age differences in self-esteem stability across 7 decades of life. We used data from four large, heterogeneous samples to address two questions: (a) How stable is self-esteem from age 14 to 83? and (b) To what extent does the stability of self-esteem vary across gender and ethnicity? The findings from this study will help to fill some of the gaps in the literature identified in Study 1, and provide a more accurate and precise picture of self-esteem stability across the life span.

Table 3
Study 2: Self-Esteem Stability Correlations for Archival Data Sets

Sample and self-esteem measure	Reliability	Stability correlation (age in years)	Disattenuated correlation (age in years)
NELS			
7-item RSE Scale	.77 (T1)	.49 (14–16)	.62 (14–16)
	.81 (T2)	.40 (14–18)	.50 (14–18)
	.83 (T3)	.55 (16–18)	.67 (16–18)
HS&B: Sophomore Cohort			
6-item RSE Scale	.69 (T1)	.47 (15–17)	.67 (15–17)
	.72 (T2)	.35 (15–21)	.48 (15–21)
	.76 (T3)	.41 (17–21)	.55 (17–21)
HS&B: Senior Cohort			
6-item RSE Scale	.70 (T1)	.46 (17–19)	.66 (17–19)
	.71 (T2)	.37 (17–23)	.51 (17–23)
	.75 (T3)	.45 (19–23)	.62 (19–23)
NLS			
4-item RSE Scale	.67 (T1)	.39 (18–19)	.59 (18–19)
	.66 (T2)	.37 (18–20)	.52 (18–20)
	.75 (T3)	.34 (18–22)	.47 (18–22)
	.77 (T4)	.48 (19–20)	.69 (19–20)
		.40 (19–22)	.56 (19–22)
		.47 (20–22)	.62 (20–22)
ACL			
3-item RSE scale			
27–30 years	.63 (T1); .63 (T2)	.48 (27–30)	.76 (27–30)
34–37 years	.61 (T1); .60 (T2)	.62 (34–37)	.99 (34–37)
44–47 years	.62 (T1); .61 (T2)	.45 (44–47)	.74 (44–47)
55–58 years	.51 (T1); .57 (T2)	.36 (55–58)	.67 (55–58)
64–67 years	.55 (T1); .61 (T2)	.46 (64–67)	.79 (64–67)
75–78 years	.51 (T1); .57 (T2)	.35 (75–78)	.65 (75–78)

Note. NELS = National Educational Longitudinal Study; RSE = Rosenberg Self-Esteem Scale; HS&B = High School and Beyond; NLS = National Longitudinal Study; ACL = Americans' Changing Lives; T1 = Time 1; T2 = Time 2; T3 = Time 3.

Method

Participants and Procedure

Data from four national samples were combined to create the database for Study 2. The database consists of a total of 74,381 (52% female) participants. The ethnic breakdown was 61% Caucasian, 15% African American, 13% Hispanic American, 3% Asian American, 1% Native American, and 7% other or did not state. The archival data sets were obtained from the Inter-University Consortium for Political and Social Research. These data have not been used in previous research on self-esteem stability and were not included in the meta-analysis reported in Study 1.[8] The four samples are described below (see also Table 3).

National Educational Longitudinal Study (NELS). NELS is a longitudinal study of students who were in eighth grade in 1988 (*M* age = 14.8). Follow-up assessments were conducted 2 and 4 years later. Students were asked about their postsecondary education participation, employment, earnings, family formation, and other activities and experiences relevant to individuals about to enter their adult lives. The initial sample included 25,000 students from across the United States. These students were selected from 1,052 public and private eighth-grade schools using a stratified, clustered national probability sample design. A dataset containing 60% of the initial sample ($N = 14{,}915$) is publicly available. The final data set contained 13,192 participants (88% of the publicly available sample) who completed the self-esteem scale in at least two assessments. The ethnic breakdown was 62% Caucasian, 13% Hispanic American, 10% African American, 6% Asian American, 1% Native American, and 8% other or did not state.

High School and Beyond (HSB). HSB is a longitudinal study of students who were high school sophomores ($M = 16.7$) and seniors ($M = 18.7$) in 1980. Follow-up assessments were conducted 2, 4, and 6 years later. The third assessment did not include self-esteem; therefore, only three waves of data were used in the present study. Students were asked about their family and religious background, perceptions of self and others, personal values, extracurricular activities, and educational expectations and aspirations. The initial sample included 58,270 high school students (30,030 sophomores and 28,240 seniors) from across the United States. These students were selected from 1,015 secondary schools using a multistage, stratified, clustered sample design; within each school, 36 sophomores and 36 seniors were randomly chosen. A subsample of these students was followed longitudinally, resulting in a final sample of 26,820 sophomores and seniors. The final data set contained 25,199 participants (94% of the longitudinal sample) who completed the self-esteem scale in at least two assessments. The ethnic breakdown was 53% Caucasian, 23% Hispanic American, 19% African American, 3% Asian American, and 2% Native American.

National Longitudinal Study (NLS). NLS is a longitudinal study of students who were high school seniors in 1972 ($M = 18.7$). Follow-up assessments were conducted 1, 2, and 4 years later. Students were asked about the kinds of factors—personal, familial, social, institutional, and

[8] Tashakkori, Thompson, Wade, and Valente (1990) reported the test–retest correlation between Time 1 and Time 2 for the senior High School and Beyond cohort.

cultural—that may affect development. The initial sample included 22,652 high school seniors from across the United States. These students were selected using a multistage, stratified probability sample design. Schools were selected with equal probabilities from 600 strata. Schools in low-income areas and schools with high proportions of minority group enrollment were sampled at twice the rate of other schools. Two schools from each of the final 600 strata were sampled and then a simple random sample of 18 students from each of the sampled schools was chosen. The final data set contained 17,995 participants (79% of the initial sample) who completed the self-esteem scale in at least two assessments. The ethnic breakdown was 70% Caucasian, 13% African American, 4% Hispanic American, 1% Asian American, 1% Native American, and 11% other or did not state.

Americans' Changing Lives (ACL). ACL is a longitudinal study of adults aged 25 to 96. The initial assessment was conducted in 1986 and a follow-up assessment was conducted 3 years later. Participants were asked about interpersonal relationships, sources and levels of satisfaction, social interactions and leisure activities, traumatic life events, perceptions of retirement, health behaviors, and utilization of health care. The initial sample included 3,617 participants from across the United States. Participants were selected using a multistage, stratified probability sample design; African Americans and those 60 years of age and over were oversampled. The final data set contained 2,801 participants aged 25 to 83 who completed the self-esteem scale in both assessments.[9] The ethnic breakdown was 67% Caucasian, 30% African American, 1% Hispanic American, 1% Asian American, and 1% Native American.

Creating aggregate databases. First, all possible test retest correlations were computed within each sample. For example, in the NELS study, participants ranged in age from 13 to 17 in the initial assessment (Time 1), and follow-up assessments were conducted 2 (Time 2) and 4 years (Time 3) later. Fifteen test–retest correlations were computed from these data: 5 from Time 1 to Time 2 for each age from 13 to 17 (time interval = 2 years); 5 from Time 1 to Time 3 for each age from 13 to 17 (time interval = 4 years); and 5 from Time 2 to Time 3 (time interval = 2 years) for each age from 15 to 19 (the age of the participants at Time 2). These correlations were then entered into one data file. Next, all possible test–retest correlations were computed separately for each demographic group. For example, in the NELS data, the 15 test–retest correlations described above were computed separately for men and women and for each ethnic group. The correlations from each demographic group were entered into separate data files. To ensure that the test–retest correlations were based on large enough samples, we eliminated all correlations that were based on sample sizes smaller than 15.

Measures

Self-esteem. In all four studies, self-esteem was assessed using an abbreviated version of the Rosenberg (1965) Self-Esteem Scale (see Table 3 for psychometric information). All studies used a 4-point Likert scale ranging from 1 (*strongly agree*) to 4 (*strongly disagree*).[10]

Demographics. Age was recorded as the assessment date minus the participant's date of birth. Gender and ethnicity were recorded from the original databases.

Time interval. The time interval associated with each test–retest correlation was recorded.

Sample size. The sample size that each test–retest correlation was based on was recorded.

Results and Discussion

Descriptive Statistics

The mean test–retest correlation, averaged across age groups, was .42 (SD = .15) and the mean disattenuated correlation was .64.[11] These values are comparable with the average level of stability found in the meta-analysis (r = .50 for the raw test–retest correlation and .64 for the disattenuated test–retest correlation).

Age Differences in Self-Esteem Stability

Figure 2 shows self-esteem test–retest correlations by age. Consistent with the results of Study 1, self-esteem stability increased from adolescence to early adulthood and decreased from middle adulthood to old age. However, Study 1 showed the peak stability occurring during the 20s, whereas Study 2 showed it peaking during the 40s. To be more precise, we computed the point in the curve where stability was maximized (following Aiken & West, 1991). This peak occurred at age 48, when the estimated disattenuated test–retest correlation was .77.

Visual inspection of Figure 2 indicates a curvilinear trend. To quantify these findings and to control for the influence of time interval, we conducted a hierarchical multiple regression analysis predicting self-esteem test–retest correlations from time interval and age modeled as a linear (age), quadratic (age^2), and cubic (age^3) function. We found that quadratic age accounted for 6% ($p < .05$) of the variance over time in self-esteem stability whereas linear age accounted for 0.0% (*ns*) of the variance and cubic age accounted for 0.7% (*ns*) of the variance. Consistent with the uncorrected correlations, quadratic age accounted for 11% ($p < .05$) of the variance in self-esteem stability, indicating that differences in scale reliability did not account for the observed age trend. Thus, the relation between self-esteem stability and age was curvilinear and best fit with a quadratic function.

Gender Differences in Self-Esteem Stability

We first tested for a main effect of gender across age groups. Men and women did not differ in stability based on the uncorrected (rs = .44 for men and .43 for women) or the disattenuated test–retest correlations (rs = .64 for men and .63 for women). We next tested whether gender moderated the relation between self-esteem stability and age. Figure 3 shows the age trajectory for the uncorrected test–retest correlations for men and women. Both genders showed a similar curvilinear pattern. The interaction between gender and quadratic age was not significant for either the uncorrected (R^2 change = 0%, *ns*) or the disattenuated test–retest

[9] The ACL data set includes a measure of cognitive functioning. Respondents were asked to answer eight common questions (e.g., today's date, mother's maiden name). Sixty-six respondents missed more than half of these questions and were dropped from the current study.

[10] The High School and Beyond study and the National Longitudinal Study included a fifth option with the anchor *no opinion*. Following Kling et al. (1999), we recorded this fifth option as 2.5 (the midpoint of the scale).

[11] To address the potential bias of the formula for disattenuating test–retest correlations, we used a multiple indicator simplex model to correct for this bias and estimate a corrected disattenuated test–retest correlation (Marsh, 1993; Marsh & Grayson, 1994). In every case, the corrected disattenuated test–retest correlations fell between the raw test–retest correlations and the disattenuated test–retest correlations computed using the traditional formula. Across the four data sets, the average magnitude of the positive bias was .08.

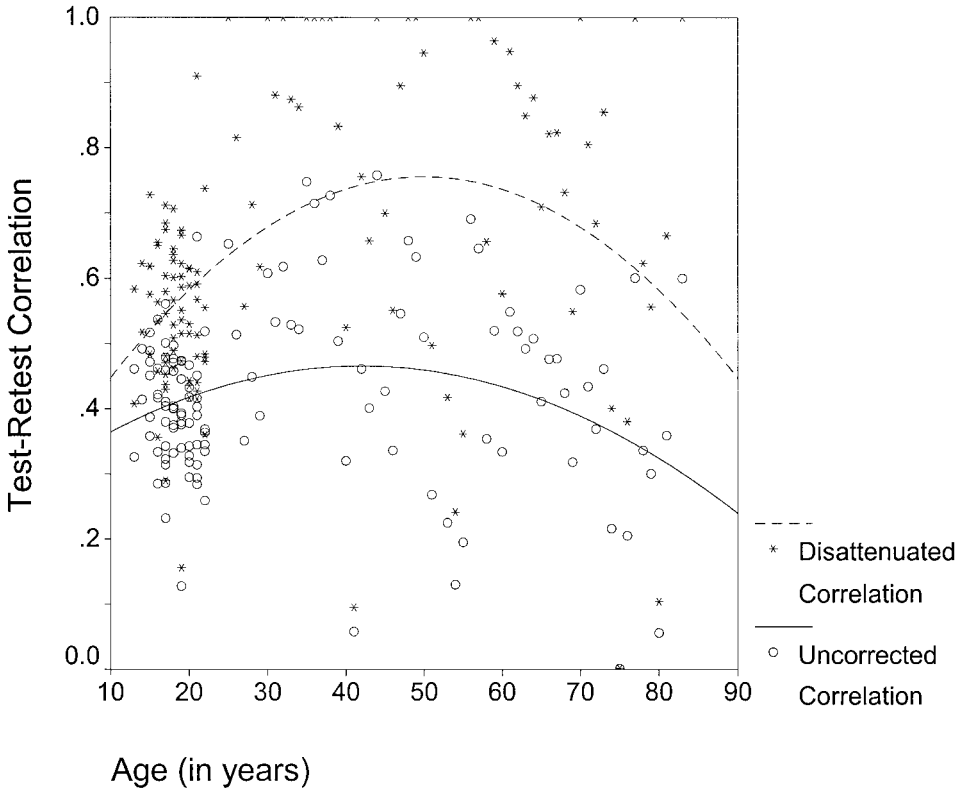

Figure 2. Study 2: Self-esteem test–retest correlations by age.

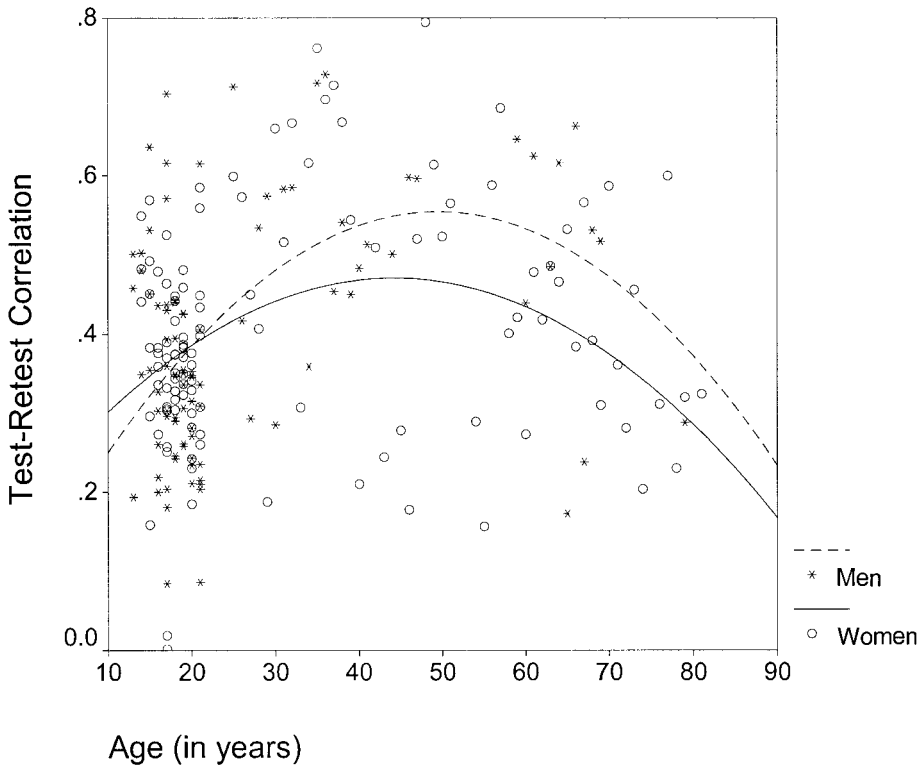

Figure 3. Study 2: Self-esteem test–retest correlations (uncorrected) by gender and age.

correlations (R^2 change = 1%, *ns*), indicating that the quadratic function fit the age trends equally well for men and women.

Ethnic Differences in Self-Esteem Stability

Asian Americans (r = .47) and Caucasians (r = .46) showed higher levels of stability than Hispanic Americans (r = .38), African Americans (r = .37), or Native Americans (r = .37), on the basis of post hoc comparisons from an analysis of variance (ANOVA) model. For the disattenuated test–retest correlations, Caucasians (r = .68) were higher than all four of the other ethnic groups (all ps < .05); Asian Americans (r = .61), African Americans (r = .60), Hispanic Americans (r = .54), and Native Americans (r = .54) did not differ significantly from each other. We next tested whether ethnicity moderated the relation between self-esteem stability and quadratic age. Less than 3% of the sample fell into the ethnic categories of Asian American, Hispanic American, or Native American beyond the mid-20s (only the ACL dataset includes participants beyond young adulthood). Therefore, the age trends could only be explored for Caucasians and African Americans. Both ethnic groups showed the curvilinear trend characterizing the total sample (see Figure 4). The interaction between ethnicity and the quadratic age was not significant for either the uncorrected (R^2 change = 1%, *ns*) or the disattenuated correlations (R^2 change = 1%, *ns*).

In summary, self-esteem stability increased from adolescence to adulthood and then decreased from adulthood to old age. This curvilinear trend was not due to age differences in scale reliability and held across gender and ethnicity. These findings are consistent with the results of Study 1.

General Discussion

The present research examined the stability of self-esteem across the life span in two studies. In Study 1, meta-analytic techniques were used to estimate self-esteem stability based on the existing research literature. In Study 2, data from four large, national studies were used to supplement the meta-analytic findings. Together, these two studies provide converging evidence about the stability of self-esteem from age 6 to 83: stability is relatively low during early childhood, increases throughout adolescence and young adulthood, and then declines during midlife and old age. This robust curvilinear trend could not be explained by age differences in the reliability of self-esteem measures and generally held across gender, ethnicity (Caucasian vs. African American), self-esteem scale, nationality (U.S. vs. non-U.S.), and year of assessment.

Several aspects of this age trajectory are noteworthy. First, we found that self-esteem was least stable during childhood. Although our analyses rule out low reliability (defined in terms of internal consistency) as an explanation for this effect, it is possible that the stabilities during childhood were attenuated by other factors. For example, self-esteem levels are generally quite elevated in childhood (Eccles, Wigfield, Harold, & Blumenfeld, 1993; Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002; Ruble, Boggiano, Feldman, & Loebl, 1980; Stipek & Tannatt, 1984) and this may

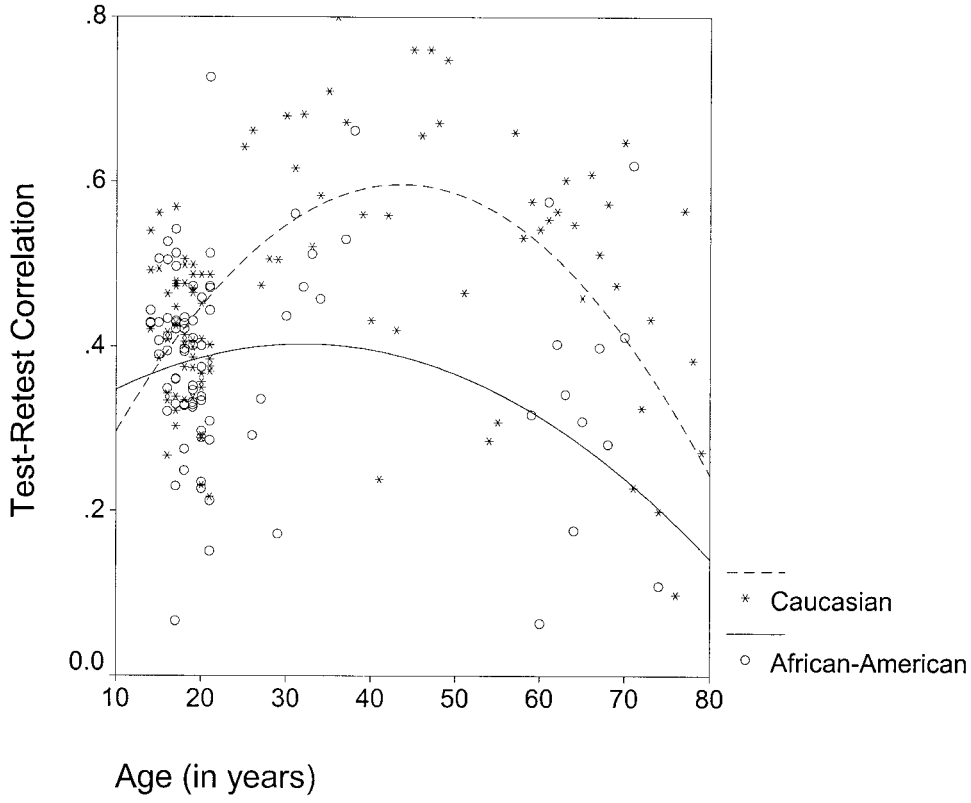

Figure 4. Study 2: Self-esteem test–retest correlations (uncorrected) by ethnicity and age.

have produced a restriction of range that would reduce test–retest correlations. Another possibility is that young children do not fully understand the meaning of the questions on self-esteem scales, but provide consistent responses based on their current mood. Relatedly, young children may lack the ability to form abstract conceptions of themselves as globally good or bad, and instead respond to self-esteem questions based on relatively transient evaluations by parents, teachers, and others. Each of these possibilities could explain the relatively low stability of self-esteem in childhood.

Researchers have debated whether self-esteem measures are valid for young children. Harter (1983; Harter & Pike, 1984) posited that children younger than 8 or 9 can evaluate themselves in specific domains, but they lack the cognitive capacity to integrate these domain-specific evaluations into a generalized, global concept of the self. As a result, she proposed that measures of global self-esteem given to young children lack construct validity (Harter, 1983). However, Marsh, Craven, and Debus (1991) questioned Harter's claim, stating that she has not provided enough convincing empirical evidence to support her position. In contrast to Harter, Marsh and his colleagues suggest that a global sense of self-worth may precede, rather than follow from, domain-specific self-representations, and therefore global self-esteem can be validly measured in young children. To support their position, Marsh et al. (1991) conducted a confirmatory factor analysis of their Self-Descriptive Questionnaire in a sample of 6- to 8-year-olds and found a reliable, well-defined global self-worth factor that was distinct from the domain-specific factors (see also Marsh, Ellis, & Craven, 2002).

This debate over the validity of self-esteem measurement in young children has important implications for the study of self-esteem stability. Quite simply, if self-esteem cannot be measured validly in early childhood, then stability and change in self-esteem cannot be assessed for this age group. Unfortunately, resolution of this issue is beyond the scope of the present research. More studies evaluating the validity of measures of global self-esteem in young children are needed to clarify whether there can be a meaningful discussion of stability and change in self-esteem during early childhood.

A second noteworthy aspect of the age trajectory is that self-esteem stability increased throughout adolescence and into early adulthood. Why is stability lower during early adolescence than during late adolescence and early adulthood? Lower stability is expected when the individual is faced with dramatic environmental and/or maturational changes (Alsaker & Olweus, 1992; B. W. Roberts & DelVecchio, 2000). Indeed, the transition to adolescence is perhaps the most volatile transition one experiences in a lifetime, entailing numerous social, cognitive, and biological changes. It is associated with rapid maturational changes, increasing cognitive abilities, shifting societal demands, exploration of new identities and roles, and the initiation of romantic and sexual relationships. These changes may impact individuals differently. For example, adolescents who physically mature earlier than their peers may lose self-esteem whereas adolescents who mature at the same time as their peers may gain self-esteem. This scenario would have the effect of shifting the relative ordering of individual differences in self-esteem, which would reduce stability over time.

Although late adolescence and early adulthood also involve numerous changes (albeit in fewer domains than in early adolescence), the individual is gradually obtaining the psychological resources and autonomy needed to adjust to these changes. For example, Erikson (1968) stated that late adolescence is a period of identity achievement and consolidation. These identities, once achieved, provide stability and help the individual navigate through the many changes that occur over the course of life. Furthermore, as individuals make the transition into adulthood, maturational changes are reduced, environmental changes are increasingly subject to individual control, and a more stable sense of self is formed. Thus, it is not surprising to find that self-esteem stability increases from early adolescence to young adulthood. These same developmental changes no doubt also contribute to the increasing stability of personality traits from adolescence to adulthood (B. W. Roberts & DelVecchio, 2000).

A third aspect of the age trajectory is that self-esteem stability decreased from adulthood to old age. Throughout midlife, people tend to experience relatively few environmental changes—social roles are generally established, identity is fairly consolidated, and career and relationship choices generally have been made. In contrast, dramatic life changes and shifting social circumstances may be more characteristic of later adulthood and old age. For example, normative life events, such as children moving out of the home, retirement, and death of a loved one, may lead to changes in social roles and corresponding shifts in identity. In addition, maturational changes again become common, such as health problems, which can result in greater dependency on others and reduced feelings of personal agency. These events may challenge some individuals' view of themselves and thus produce idiosyncratic changes that reduce the stability of self-esteem.

Another possibility is that as individuals age they may begin to review their lifelong accomplishments and experiences, leading in some cases to more critical self-appraisals and in other cases to greater acceptance of their faults and limitations. This process would be consistent with Erikson's (1985) notion of old age as a time when people reflect on the life they have lived. Some individuals may decide that they have succeeded in life (i.e., develop ego integrity) and thus maintain or increase their self-esteem whereas others may decide that they have failed (i.e., suffer despair) and experience a decline in self-esteem. Thus, a developmental shift toward greater self-reflection in old age may produce increases in self-esteem for some individuals but decreases for others. This process would contribute to lower levels of stability.

Of interest, at no time did the stability of self-esteem reach unity. Self-esteem stability (disattenuated test–retest correlations) peaked at .71 in Study 1 and .77 in Study 2. Thus, self-esteem continues to change across the life span. The evidence suggests that self-esteem, much like personality traits, is never set like plaster and, in fact, shows even lower levels of stability in old age. Nonetheless, we did find substantial levels of continuity across decades of life, suggesting that self-esteem is best characterized as showing both continuity and change across the life span.

What are the implications of these findings for self-esteem's position in the hierarchy of consistency (e.g., Conley, 1984; Kelly, 1955)? B. W. Roberts and DelVecchio's (2000) meta-analysis of personality trait stability provides a useful comparison point. In general, self-esteem and personality traits show similar levels of stability. Roberts and DelVecchio reported uncorrected personality test–retest correlations (controlling for time interval) of .43 in childhood, .44 in adolescence, .54 during the college years, .60 for ages 22 to 29, and .64 during the 30s. Similarly, in our meta-

analysis (Study 1), we found uncorrected correlations (controlling for time interval) of .40 in childhood, .48 in adolescence, .55 during the college years, .65 for ages 22 to 29, and .62 during the 30s. However, B. W. Roberts and DelVecchio reported a test–retest correlation of .74 for adults between the ages of 50 and 70 whereas we found a test–retest correlation of .49 for this same age range.

Considering the similarity in the stability of self-esteem and personality throughout most of the life span, this large divergence during late adulthood seems puzzling. It is possible that the developmental shift toward greater self-reflection described above produces changes only in self-esteem, without affecting basic personality traits. Indeed, reflecting on the overall worthiness of your life may change your self-esteem but it is not likely to have a powerful impact on how sociable, conscientious, and creative you are. To test this speculative account, more research is needed to understand the self-evaluative processes that underline continuity and change in self-esteem over the life span.

In general, our results are not consistent with Conley's (1984) conclusions about the hierarchy of consistency. Conley argued that personality traits and self-esteem are "distinct orders of phenomena characterized by different levels of longitudinal consistency in adulthood" (p. 11). He reported that self-esteem, in contrast to personality traits, is not longitudinally stable across adolescence, adulthood, or old age. Contrary to this conclusion, our results indicate that self-esteem was as stable as personality traits throughout much of the life span, diverging only in old age.

These findings have implications for the state-trait distinction. If self-esteem is just as stable as personality, then it should be conceptualized as a trait-like construct. How can we reconcile these findings with experimental research showing that self-esteem levels can be increased or decreased through laboratory manipulations? In our view, experiments that produce transient changes in self-evaluations do not support the conclusion that self-esteem is merely a state. Just as situational factors can make introverts act outgoing, disagreeable people act kind, and emotionally stable people feel tense and stressed, situations created in the lab can produce temporary changes in self-esteem. To say that self-esteem is a trait does not imply that it is impervious to experience. As Gordon Allport (1966) noted, "I do not perspire except in the heat, nor shiver except in the cold" but there may nonetheless be stable individual differences in the tendency to shiver and to perspire (p. 2). Similarly, although experimental manipulations can produce temporary drops in self-evaluations, such evidence is not inconsistent with the notion that self-esteem is a stable trait. The only way to determine whether a construct is a trait is to examine its stability over long periods of time.

The present findings, which suggest that people maintain a similar relative level of self-esteem despite experiencing the inevitable successes and failures of life, are consistent with a trait view of self-esteem. Nonetheless, we feel that there is merit to conceptualizing self-esteem as both state and trait. To characterize self-esteem as entirely trait-like may obscure the fact that changes can and do occur in response to various experiences and interventions. Similarly, to characterize self-esteem as entirely state-like obscures the degree to which self-esteem is consistent over time, linked to highly stable individual-difference constructs such as personality traits, predictive of long-term life outcomes, and moderately heritable (e.g., Kendler, Gardner, & Prescott, 1998; McGuire et al., 1999; Roy, Neale, & Kendler, 1995).

Finally, our findings have implications for the timing of interventions aimed at changing self-esteem. Previous research suggests that programs designed to increase the self-esteem of children and adolescents can be successful and have beneficial effects for a wide range of developmental outcomes (e.g., DuBois, in press; Haney & Durlak, 1998). Our results suggest that interventions should target early adolescence because self-esteem stability is relatively low and there is little controversy over the validity of self-reports in that age group. If relatively more "true" change is taking place during early adolescence, it may offer an important window of opportunity for programs designed to influence self-esteem levels.

The present research has several limitations that should be considered when interpreting the findings. First, both studies suffer from sampling problems. Study 1, like all meta-analyses, is plagued by the well-known file drawer problem (Rosenthal, 1979). However, this concern is ameliorated in the present context because it is not likely that a researcher would disregard a study because of low test–retest correlations, particularly when most of the studies included in the meta-analysis focused on issues other than the stability of self-esteem. Generalizability is also a concern because the studies included in Study 1 were rarely based on representative samples, and in many cases used highly selective samples of convenience. Study 2 improves on the existing literature by using more representative sampling techniques, but the ACL dataset, which provided the majority of the correlations for Study 2, is not fully representative. Although the ACL study is a national study and is heterogeneous in terms of socioeconomic status, it is not representative of the U.S. population in terms of ethnicity.[12]

Second, most studies reported test–retest correlations over relatively short age spans—across both studies, the average time interval between assessments was 3 years. Although our analyses allow us to estimate the level of stability we would expect to see over much longer periods of time, research based on longer time intervals would provide a more accurate picture of self-esteem stability across the life span.

Third, neither Study 1 nor Study 2 provides information about the mechanisms underlying the observed levels of self-esteem stability. For example, although the findings suggest that self-esteem stability declines in old age, the data do not provide any insights into the specific factors or mechanisms that cause this effect. Therefore, we do not know which of the many changes that occur during old age reduce the stability of self-esteem. Future research is needed to understand the processes that produce stability and change in self-esteem across the life span.

[12] To address the oversampling of African Americans and individuals aged 60 and over, we reran the analyses presented in Study 2 using a weighting variable provided in the ACL dataset. The weighted sample corresponds to 1985 United States Bureau of Census population estimates of ethnicity (Caucasians and African Americans), gender, age (25 to 64 and 65+), and region (Northeast, Midwest, South, and West). All of the findings presented in Study 2 held in the weighted sample.

References

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage.

Allport, G. W. (1966). Traits revisited. *American Psychologist, 21,* 1–10.

Alsaker, F. D., & Olweus, D. (1992). Stability of global self-evaluations on early adolescence: A cohort longitudinal study. *Journal of Research on Adolescence, 2,* 123–145.

Antonucci, T. C., & Jackson, J. S. (1983). Physical health and self-esteem. *Family and Community Health, 6,* 1–9.

Bachman, J. G., & O'Malley, P. M. (1977). Self-esteem in young men: A longitudinal analysis of the impact of educational and occupational attainment. *Journal of Personality and Social Psychology, 35,* 365–380.

Baumeister, R. F., Smart, L., & Boden, J. M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review, 103,* 5–33.

Block, J. H., Gjerde, P. F., & Block, J. H. (1991). Personality antecedents of depressive tendencies in 18-year-olds: A prospective study. *Journal of Personality and Social Psychology, 60,* 726–738.

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.

Brown, J. D., & Mankowski, T. A. (1993). Self-esteem, mood, and self-evaluation: Changes in mood and the way you see you. *Journal of Personality and Social Psychology, 64,* 421–430.

Campbell, A. (1981). *The sense of well-being in America: Recent patterns and trends.* New York: McGraw-Hill.

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.

Conley, J. J. (1984). 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.

Coopersmith, S. (1967). *The antecedents of self-esteem.* San Francisco: Freeman.

Cruz-Perez, R. (1973). The effect of experimentally induced failure, self-esteem, and sex on cognitive differentiation. *Journal of Abnormal Psychology, 81,* 74–79.

Demo, D. H. (1992). The self-concept over time: Research issues and directions. *Annual Review of Sociology, 18,* 303–326.

DeNeve, K. M., & Cooper, H. (1998). The happy personality: A meta-analysis of 137 personality traits and subjective well-being. *Psychological Bulletin, 124,* 197–229.

Diener, E., & Diener, M. (1995). Cross-cultural correlates of life satisfaction and self-esteem. *Journal of Personality and Social Psychology, 68,* 653–663.

Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., Moffitt, T. E., & Caspi, A. (2002). *Exploring the link between self-esteem and externalizing problems: Low self-esteem is related to antisocial behavior and delinquency.* Manuscript under review. University of California, Davis, Department of Human Development.

DuBois, D. L. (in press). Promoting the self-esteem of children. In T. P. Gullotta & M. Bloom (Eds.) & L. Bond (Section Ed.), *Encyclopedia of primary prevention and health promotion: Childhood*. New York: Kluwer Academic/Plenum.

Dumenci, L., & Windle, M. (1996). The trait-state model of adolescent depression using the Center for Epidemiology Studies Depression Scale. *Multivariate Behavioral Research, 31,* 313–330.

Eccles, J., Wigfield, A., Harold, R. D., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64,* 830–847.

Elliott, M. (1996). Impact of work, family, and welfare receipt on women's self-esteem in young adulthood. *Social Psychology Quarterly, 59,* 80–95.

Erikson, E. H. (1968). *Identity, youth, and crisis* (1st ed.) New York: Norton.

Erikson, E. H. (1985). *The life cycle completed: A review.* New York: Norton.

Fordyce, M. W. (1988). A review of research on the happiness measure: A sixty second index of happiness and mental health. *Social Indicators Research, 20,* 355–382.

Gergen, K. J. (1981). The functions and foibles of negotiating self-conceptions. In M. D. Lynch, A. A. Norem-Hebeisen, & K. J. Gergen (Eds.), *Self-concept: Advances in theory and research* (pp. 59–73). Cambridge, MA: Ballinger.

Granleese, J., & Joseph, S. (1994). Reliability of the Harter Self-Perception Profile for Children and predictors of global self-worth. *Journal of Genetic Psychology, 155,* 487–492.

Greenberg, J., Solomon, S., Pyszczynski, T., Rosenblatt, A., Burling, J., Lyon, D., et al. (1999). Why do people need self-esteem? Converging evidence that self-esteem serves as an anxiety-buffering function. In R. F. Baumeister (Ed.), *The self in social psychology* (pp. 105–118). Philadelphia: Taylor & Francis.

Haney, P., & Durlak, J. A. (1998). Changing self-esteem in children and adolescents: A meta-analytic review. *Journal of Clinical Child Psychology, 27,* 423–433.

Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research, 52,* 123–142.

Harter, S. (1982). The perceived competence scale for children. *Child Development, 53,* 87–97.

Harter, S. (1983). Developmental perspectives on the self-system. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology* (pp. 276–385). New York: Wiley.

Harter, S. (1985). *Manual for the Self-Perception Profile for Children (revision of the Perceived Competence Scale for Children).* Denver, CO: University of Denver.

Harter, S. (1998). The development of self-representations. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology* (pp. 553–617). New York: Wiley.

Harter, S., & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child Development, 55,* 1969–1982.

Heatherton, T. F., Herman, C. P., & Polivy, J. (1991). Effects of physical threat and ego threat on eating behavior. *Journal of Personality and Social Psychology, 60,* 138–143.

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology, 60,* 895–910.

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

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

Hendrick, S. S., Hendrick, C., & Adler, N. L. (1988). Romantic relationships: Love, satisfaction, and staying together. *Journal of Personality and Social Psychology, 54,* 980–988.

Hirsch, B. J., & DuBois, D. L. (1991). Self-esteem in early adolescence: The identification and prediction of contrasting longitudinal trajectories. *Journal of Youth and Adolescence, 20,* 53–72.

Jones, E. E., Rhodewalt, F., Berglas, S., & Skelton, J. A. (1981). Effects of strategic self-presentation on subsequent self-esteem. *Journal of Personality and Social Psychology, 41,* 407–421.

Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emo-

tional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology, 86,* 80–92.

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

Kendler, K. S., Gardner, C. O., & Prescott, C. A. (1998). A population-based twin study of self-esteem and gender. *Psychological Medicine, 28,* 1403–1409.

Kenny, D. A., & Zautra, A. (2001). Trait-state models for longitudinal data. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 243–263). Washington, DC: American Psychological Association.

Kernis, M. H., Cornell, D. P., Sun, C., Berry, A., & Harlow, T. (1993). There's more to self-esteem than whether it is high or low: The importance of stability of self-esteem. *Journal of Personality and Social Psychology, 65,* 1190–1204.

Kernis, M. H., Grannemann, B. D., & Barclay, L. C. (1992). Stability of self-esteem: Assessment, correlates, and excuse making. *Journal of Personality, 60,* 621–644.

Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin, 125,* 470–500.

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–62). San Diego, CA: Academic Press.

Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an interpersonal monitor: The sociometer hypothesis. *Journal of Personality and Social Psychology, 68,* 518–530.

Liu, X., Kaplan, H. B., & Risser, W. (1992). Decomposing the reciprocal relationships between academic achievement and general self-esteem. *Youth & Society, 24,* 123–148.

Major, B., Barr, L., Zubek, J., & Babey, S. H. (1999). Gender and self-esteem: A meta-analysis. In W. Swann & J. Langlois, (Eds.), *Sexism and stereotypes in modern society: The gender science of Janet Taylor Spence* (pp. 223–253). Washington DC: American Psychological Association.

Markus, H., & Kunda, Z. (1986). Stability and malleability of the self-concept. *Journal of Personality and Social Psychology, 51,* 858–866.

Marsh, H. W. (1993). Stability of individual differences in multiwave panel studies: Comparison of simplex models and one-factor models. *Journal of Educational Measurement, 30,* 157–183.

Marsh, H. W., Craven, R. G., & Debus, R. L. (1991). Self-concepts of young children 5 to 8 years of age: Measurement and multidimensional structure. *Journal of Educational Psychology, 83,* 377–392.

Marsh, H. W., Craven, R. G., & Debus, R. L. (1998). Structure, stability, and development of young children's self-concepts: A multicohort-multioccasion study. *Child Development, 69,* 1030–1053.

Marsh, H. W., Ellis, L. A., & Craven, R. G. (2002). How do preschool children feel about themselves? Unraveling measurement and multidimensional self-concept structure. *Developmental Psychology, 38,* 376–393.

Marsh, H. W., & Grayson, D. (1994). Longitudinal stability of latent means and individual differences: A unified approach. *Structural Equation Modeling, 1,* 317–359.

Marsh, H. W., Smith, I. D., & Barnes, J. (1983). Multitrait-multimethod analyses of the Self-Description Questionnaire: Student–teacher agreement on multidimensional ratings of student self-concept. *American Educational Research Journal, 20,* 333–357.

McFarlin, D. B., Baumeister, R. F., & Blascovich, J. (1984). On knowing when to quit: Task failure, self-esteem, advice, and nonproductive persistence. *Journal of Personality, 52,* 138–155.

McGuire, S., Manke, B., Saudino, K., Reiss, D., Hetherington, E. M., & Plomin, R. (1999). Perceived competence and self-worth during adolescence: A longitudinal behavioral genetic study. *Child Development, 70,* 1283–1296.

Mortimer, J. T., Finch, M. D., & Kumba, D. (1982). Persistence and change in development: The multidimensional self-concept. *Life-Span Development and Behavior, 4,* 263–313.

Murray, S. L., Holmes, J. G., & Griffin, D. W. (2000). Self-esteem and the quest for felt security: How perceived regard regulates attachment processes. *Journal of Personality and Social Psychology, 78,* 478–498.

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.

Musser, L. M., & Browne, B. A. (1991). Self-monitoring in middle childhood: Personality and social correlates. *Developmental Psychology, 27,* 994–999.

Neyer, F. J., & Asendorpf, J. B. (2001). Personality-relationship transaction in young adulthood. *Journal of Personality and Social Psychology, 81,* 1190–1204.

Nottelmann, E. D. (1987). Competence and self-esteem during transition from childhood to adolescence. *Developmental Psychology, 23,* 441–450.

O'Connor, B. P., & Vallerand, R. J. (1998). Psychological adjustment variables as predictors of mortality among nursing home residents. *Psychology and Aging, 13,* 368–374.

Owens, T. J. (1994). Two dimensions of self-esteem: Reciprocal effects of positive self-worth and self-deprecation on adolescent problems. *American Sociological Review, 59,* 391–407.

Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology, 74,* 1197–1208.

Reinherz, H. Z., Giaconia, R. M., Pakiz, B., & Silverman, A. B. (1993). Psychosocial risks for major depression in late adolescence: A longitudinal community study. *Journal of the American Academy of Child and Adolescent Psychiatry, 32,* 1155–1163.

Reitzes, D. C., Mutran, E. J., & Fernandez, M. E. (1996). Does retirement hurt well-being? Factors influencing self-esteem and depression among retirees and workers. *The Gerontologist, 36,* 649–656.

Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin, 126,* 3–25.

Roberts, J. E., Gotlib, I. H., & Kassel, J. D. (1996). Adult attachment security and symptoms of depression: The mediating roles of dysfunctional attitudes and low self-esteem. *Journal of Personality and Social Psychology, 70,* 310–320.

Roberts, R. E., & Bengston, V. L. (1996). Affective ties to parents in early adulthood and self-esteem across 20 years. *Social Psychology Quarterly, 59,* 96–106.

Robins, R. W., Fraley, R. C., Roberts, B. W., & Trzesniewski, K. H. (2001). A longitudinal study of personality change in young adulthood. *Journal of Personality, 69,* 617–640.

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single item measure and the Rosenberg Self-Esteem scale. *Personality and Social Psychology Bulletin, 27,* 151–161.

Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Self-esteem across the lifespan. *Psychology and Aging, 17,* 423–434.

Roeser, R. W., & Eccles, J. S. (1998). Adolescents' perceptions of middle school: Relation to longitudinal changes in academic and psychological adjustment. *Journal of Research on Adolescence, 8,* 123–158.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Rosenberg, M., Schooler, C., & Schoenbach, C. (1989). Self-esteem and

adolescent problems: Modeling reciprocal effects. *American Sociological Review, 54,* 1004–1018.

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review, 60,* 141–156.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86,* 638–641.

Roy, M. A., Neale, M. C., & Kendler, K. S. (1995). The genetic epidemiology of self-esteem. *British Journal of Psychiatry, 166,* 813–820.

Rubin, R. A. (1978). Stability of self-esteem ratings and their relation to academic achievement: A longitudinal study. *Psychology in the Schools, 15,* 430–433.

Ruble, D. N., Boggiano, A. K., Feldman, N. S., & Loebl, J. H. (1980). Developmental analysis of the role of social comparison in self-evaluation. *Developmental Psychology, 16,* 105–115.

Shrauger, J. S., & Sorman, P. B. (1977). Self-evaluations, initial success and failure, and improvement as determinants of persistence. *Journal of Consulting and Clinical Psychology, 45,* 784–795.

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.

Stipek, D. J., & Tannatt, L. M. (1984). Children's judgments of their own and their peers' academic competence. *Journal of Educational Psychology, 76,* 75–84.

Tashakkori, A., Thompson, V. D., Wade, T. J., & Valente, E. (1990). Structure and stability of self-esteem in late teens. *Personality and Individual Differences, 11,* 885–893.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology, Vol. 21* (pp. 181–227). San Diego, CA: Academic Press.

Vingilis, E., Wade, T. J., & Adlaf, E. (1998). What factors predict student self-rated physical health? *Journal of Adolescence, 21,* 83–97.

Whisman, M. A., & Kwon, P. (1993). Life stress and dysphoria: The role of self-esteem and hopelessness. *Journal of Personality and Social Psychology, 65,* 1054–1060.

Wilson, W. (1967). Correlates of avowed happiness. *Psychological Bulletin, 67,* 294–306.

Wylie, R. C. (1979). *The self-concept.* Lincoln: University of Nebraska Press.

Received April 15, 2001
Revision received July 12, 2002
Accepted July 16, 2002 ■