Self-Esteem Development Across the Life Span: A Longitudinal Study With a Large Sample From Germany

Ulrich Orth
University of Bern

Jürgen Maes
Bundeswehr University Munich

Manfred Schmitt
University of Koblenz–Landau

The authors examined the development of self-esteem across the life span. Data came from a German longitudinal study with 3 assessments across 4 years of a sample of 2,509 individuals ages 14 to 89 years. The self-esteem measure used showed strong measurement invariance across assessments and birth cohorts. Latent growth curve analyses indicated that self-esteem follows a quadratic trajectory across the life span, increasing during adolescence, young adulthood, and middle adulthood, reaching a peak at age 60 years, and then declining in old age. No cohort effects on average levels of self-esteem or on the shape of the trajectory were found. Moreover, the trajectory did not differ across gender, level of education, or for individuals who had lived continuously in West versus East Germany (i.e., the 2 parts of Germany that had been separate states from 1949 to 1990). However, the results suggested that employment status, household income, and satisfaction in the domains of work, relationships, and health contribute to a more positive life span trajectory of self-esteem. The findings have significant implications, because they call attention to developmental stages in which individuals may be vulnerable because of low self-esteem (such as adolescence and old age) and to factors that predict successful versus problematic developmental trajectories.

Keywords: self-esteem, life span development, measurement invariance, longitudinal

Supplemental materials: http://dx.doi.org/10.1037/a0038481.supp

How the self-esteem of individuals develops as they go through life is of considerable societal significance, because mounting evidence suggests that self-esteem is predictive of a person’s success and well-being in important life domains such as relationships, work, and health. For example, longitudinal research indicates that self-esteem predicts satisfaction in marriage and close relationships (Orth, Robins, & Widaman, 2012), social support (Marshall, Parker, Ciarrochi, & Heaven, 2014), success and well-being in working life (Kuster, Orth, & Meier, 2013; Orth et al., 2012), mental health (Sowislo & Orth, 2013; Trzesniewski et al., 2006), and physical health (Orth et al., 2012; Trzesniewski et al., 2006). In recent years, a growing number of studies have provided important insights into the nature and consequences of self-esteem development (for a review, see Orth & Robins, 2014). However, few longitudinal studies have examined self-esteem from a life-span perspective because few have included samples with sufficiently broad age ranges. As yet, two longitudinal data sets have been used to track the self-esteem trajectory across the life span (Orth et al., 2012; Orth, Trzesniewski, & Robins, 2010). However, though these studies yielded converging evidence on the life-span trajectory, their samples came from a single country (the United States). Given that cultures shape the prototypical self-concept and tendency toward self-enhancement among their members (Heine, Lehman, Markus, & Kitayama, 1999; Markus & Kitayama, 1991), thus potentially affecting the normative self-esteem trajectory, it is important to examine the life-span trajectory in samples from other cultural contexts.

In this article, we use data from a large longitudinal study from Germany that included individuals ages 14–89 years to (a) study the shape of the life-span trajectory of self-esteem, (b) test for generational changes in the trajectory, and (c) examine moderating factors of self-esteem development. Moreover, given that the study was conducted in the 1990s, a few years after German reunification in 1990, and given that people in West versus East Germany had long (from 1949 to 1990) been living in two separate states with important political, economic, and social differences, we used the unique design of the study to test whether the findings replicated across these two cultural contexts within Germany.

Before describing the current state of knowledge on the development of self-esteem, we briefly define the construct. Self-esteem refers to an “individual’s subjective evaluation of his or her worth as a person” (Donnellan, Trzesniewski, & Robins, 2011, p. 718; see also Harter, 2006). It is important to note that self-esteem does not necessarily reflect an individual’s objective abilities or social
status. Moreover, self-esteem is typically defined as the “feeling that one is ‘good enough,’” although high self-esteem does not necessarily imply that a person considers him- or herself superior to others (Rosenberg, 1965, p. 31). Thus, self-esteem involves feelings of self-acceptance and self-respect, but the construct is conceptually and empirically distinct from narcissism (Ackerman et al., 2011). Empirical evidence shows that individual differences in self-esteem are relatively stable across time (Trzesniewski, Donnellan, & Robins, 2003), even across several decades (Kuster & Orth, 2013), suggesting that the stability of self-esteem is similar to the stability of the Big Five personality factors (Roberts & DelVecchio, 2000). Hence, self-esteem should be considered a personality trait.

Theoretical Perspectives on Life-Span Development of Self-Esteem

Although research indicates that self-esteem changes systematically as people age, no theory is yet available that focuses specifically on the development of self-esteem across the life span. Within the broader context of personality development, a number of theories have been suggested (for a review, see Specht et al., 2014), of which neo-socioanalytic theory (Roberts & Wood, 2006; see also Roberts, Wood, & Caspi, 2008) is of particular relevance for deriving hypotheses about how and why self-esteem changes across the life course. Although the neo-socioanalytic theory accounts for genetic influences on personality, a central proposition is that environmental factors can influence personality traits at any age (a proposition called the plasticity principle). In particular, the theory emphasizes the importance of social roles for personality development, stating that social roles are the primary channel through which the environment influences personality. People assume social roles in life domains such as close relationships, family, work, and the community, and each social role involves a set of social expectations that describe which behavior is role congruent and, consequently, which behavior will be rewarded or punished by others who interact with an individual. These expectations influence behavior and, in the long run, may lead to changes in an individual’s personality. According to neo-socioanalytic theory, many social roles (such as the roles of spouse, parent, employee, and supervisor) call for emotional stability, conscientiousness, social dominance, and agreeableness.

During young adulthood, people establish many new social roles—for example, by entering into working life, by marrying (or cohabiting with a relationship partner), or by having a baby (Hutteman, Hennecke, Orth, Reitz, & Specht, 2014). During middle adulthood, developmental tasks typically consist in further investing in these social roles by maintaining a satisfying relationship with one’s spouse or relationship partner; by helping children to become responsible and well-functioning adults; and by further improving one’s professional competencies and, frequently, taking on more managerial functions at work (Hutteman et al., 2014). In both young and middle adulthood, most people invest heavily in their social roles, and—according to neo-socioanalytic theory—therefore develop in the direction of personality characteristics that help them to function better in these social roles (as described by the social investment principle and the maturity principle; Roberts & Wood, 2006; Roberts et al., 2008). Mature personality traits such as emotional stability, conscientiousness, social dominance, and agreeableness are associated with self-esteem (Robins, Hendin, & Trzesniewski, 2001; Watson, Suls, & Haig, 2002). Moreover, self-esteem is itself a characteristic that improves functioning in social roles (Orth et al., 2012). Consequently, neo-socioanalytic theory suggests that self-esteem shows normative increase during young and middle adulthood.

With regard to old age, however, theoretical perspectives suggest that self-esteem might no longer show normative increase but, rather, decrease (Orth et al., 2010). First, the transition to old age frequently includes negative changes in social roles, in particular due to retirement from work; the empty nest situation (i.e., when a person’s children have grown up and left home); and, possibly, widowhood. Consistent with the emphasis on social roles in neo-socioanalytic theory, empirical studies tend to show normative decrease in conscientiousness during old age (Lucas & Donnellan, 2011; Marsh, Nagengast, & Morin, 2013; Specht, Egloff, & Schmukle, 2011), which might entail a decrease in self-esteem. Second, transitions in old age not only affect social roles but may also lead to negative changes in social relationships. For example, retirement significantly reduces or even ends contact with former coworkers, and widowhood may—beside the loss of the relationship partner itself—negatively alter one’s relationships with children and one’s spouse’s family and friends. Given that relationships are one of the most influential factors in self-esteem (Leary & Baumeister, 2000; MacDonald & Leary, 2012), significant loss in the relationship domain might lead to reduced self-esteem in old age. Third, in old age, people often experience negative changes with regard to other important sources of self-esteem such as socioeconomic status (as indicated by income and occupational prestige) and health (e.g., reduced mobility and declining cognitive and physical abilities; Baltes & Mayer, 1999; Wagner, Gerstorf, Hoppmann, & Luszcz, 2013), which might further impair self-esteem.

Previous Research on Life-Span Development of Self-Esteem

Empirical evidence is provided by two longitudinal studies that have examined the life-span trajectory of self-esteem (Orth et al., 2010, 2012). In both studies, a curvilinear trajectory with an inverted U shape provided the best fit to the data: Self-esteem increased from adolescence to middle adulthood, peaked in the range of age 50–60 years, and then decreased into old age. Across studies, the increase corresponded to about ½ to ⅔ of a standard deviation and the decrease to about ⅔ of a standard deviation. Additional longitudinal studies have focused on specific developmental periods such as adolescence (Birkeland, Melkevik, Holsen, & Wold, 2012; Erol & Orth, 2011; Steiger, Allemand, Robins, & Fend, 2014), young adulthood (Chung et al., 2014; Wagner, Lang, Neyer, & Wagner, 2014; Wagner, Lüdtke, Jonkmann, & Trautwein, 2013; Zeiders, Umaña-Taylor, & Derlan, 2013), and old age (Wagner, Gerstorf, et al., 2013; Wagner et al., 2014). Overall, the results of these studies are consistent with the life-span trajectory summarized earlier. Further evidence is provided by a meta-analysis of mean-level change in self-esteem, which suggests that self-esteem increases in young adulthood but does not change after age 30 years; however, few studies were available after age 30,
limiting the statistical power of analyses in middle and older adulthood (Huang, 2010).

In addition, cross-sectional studies have tested for age differences in self-esteem over the life course (McMullin & Cairney, 2004; Pullmann, Allik, & Realo, 2009; Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002). Whereas the findings of Robins et al. (2002) are relatively similar to the longitudinal evidence reviewed earlier, the results of McMullin and Cairney (2004) and Pullmann et al. (2009) diverge more strongly. A possible explanation for this inconsistency is that cross-sectional data generally suffer from important limitations, such as confounding age-related changes with cohort effects (Baltes, Cornelius, & Nesselroade, 1979). To overcome this problem, cohort-sequential data are required—that is, longitudinal data from multiple birth cohorts. Then, growth curve analyses allow testing whether a single coherent trajectory can be modeled across cohorts or whether cohort membership accounts for systematic differences in the growth factors of the life-span trajectory (Duncan, Duncan, & Stryer, 2006; Preacher, Wichman, MacCallum, & Briggs, 2008).

In fact, there is an ongoing debate about whether there have been generational changes in self-esteem development during the past decades (Gentile, Twenge, & Campbell, 2010; Trzesniewski & Donnellan, 2009, 2010; Twenge & Campbell, 2001). On the one hand, it is possible that sociocultural changes have influenced the self-esteem trajectory of recent generations (including both the level and shape of the trajectory). For example, secular changes such as an increasing cultural focus on self-esteem (in families, schools, workplaces, and the media) and the advent of social media, with increasing possibilities for self-presentation, might lead to a higher level or a steeper increase in self-esteem. On the other hand, findings from cohort-sequential studies do not support the claim of generational changes in the self-esteem trajectory, either in the overall level or in the shape of the trajectory (Erol & Orth, 2011; Orth et al., 2010, 2012). Instead, the evidence suggests that the average life-span trajectory has not changed over the last century. Nevertheless, in the present research, we tested for generational changes in self-esteem, given that previous cohort-sequential studies were based exclusively on data from the United States.

Previous research has tested for factors that explain individual differences in the life-span trajectory of self-esteem. Although, on average, self-esteem changes according to the normative trajectory described earlier, individuals differ substantially in the particular trajectory they follow—interindividual variability that is captured by the variances of growth factors in latent growth curve models. For example, in the United States, ethnic minorities did not show the same self-esteem trajectory as majority group members; whereas the trajectory of Black people (compared to that of White people) was more positive during young adulthood (Erol & Orth, 2011), their self-esteem declined much more sharply in older adulthood (Orth et al., 2010). Moreover, individuals with higher levels of education reported higher self-esteem across the life span than did individuals with lower levels of education (Orth et al., 2010). Also, as suggested by the theoretical perspectives outlined earlier, empirical evidence indicates that success in the domains of relationships, work, and health moderates the life-span self-esteem trajectory (Orth et al., 2010; Wagner, Lüdtke et al., 2013). In addition, the available data suggest that emotional stability, conscientiousness, and extraversion predict more positive self-esteem trajectories (Erol & Orth, 2011; Wagner, Lüdtke et al., 2013).

The Present Research

The first goal of this research was to examine the life-span trajectory of self-esteem, using longitudinal data from a large sample from Germany. We tested which of several growth models (intercept only, linear, quadratic, and cubic) yielded the best fit to the data. Given that previous studies on the life-span trajectory were based on samples from the United States, the present study allows for evaluation of whether the U.S. findings replicate in a different cultural context. The second goal was to use the cohort-sequential design of the present research to test whether there is evidence of generational changes in the life-span self-esteem trajectory. The third goal was to test for moderators of the self-esteem trajectory. Specifically, we tested for the effects of demographic variables (gender, education, residence in West vs. East Germany) and covariates in the domains of relationships (relationship status, relationship satisfaction), work (employment status, job satisfaction), wealth (individual income, household income), and physical health (health satisfaction).

Moreover, we had two methodological goals. First, we tested whether the self-esteem measure used showed measurement invariance across cohorts and time. Measurement invariance is essential for growth modeling because scores for different cohorts and scores at different waves are comparable only when measurement invariance holds (Edwards & Wirth, 2009; Widaman, Ferrer, & Conger, 2010). Thus, results of growth curve analyses are valid only under the assumption of measurement invariance—an assumption that is, however, rarely tested in research on self-esteem development. Therefore, in a series of increasingly restricted measurement models, we tested for measurement invariance across time (i.e., three waves of data) and across birth cohorts. A second methodological goal was to build on the analyses of measurement models and to examine self-esteem as a latent construct. The crucial advantage of using latent, instead of observed, indicators in growth curve analyses is that the biasing influence of measurement error (which may be substantial) is controlled for (Cole & Preacher, 2014).

Method

The data come from the study Gerechtigkeit als innerdeutsches Problem (GIP [Justice in united Germany]; Schmitt & Maes, 1998, 2002), a longitudinal study designed to examine the psychological consequences of German reunification in the Western and Eastern parts of Germany (i.e., the two parts of Germany that had been separate states from 1949 to 1990). The main focus of the study was on the quality of life after reunification in 1990. Participants were assessed in 1996, 1998, and 2000. Random samples of participants were recruited on the basis of a geographical division of Germany into 18 cells by crossing the three factors (a) West versus East; (b) north, center, and south; and (c) small, medium-sized, and large communities. For each cell, resident registration offices provided probability samples from the population of all inhabitants between 15 and 75 years of age. Additional respondents were randomly drawn from telephone directories. Overall, about 32,000 individuals were invited to participate in the study. A
total of 3,170 individuals (i.e., 10%) agreed to participate in the study and were sent questionnaires; of these, 2,531 (i.e., 80% of those who agreed) participated in the study. For the present research, we excluded 22 participants whose age was unknown or who did not provide data on self-esteem at any of the three assessments.

Participants

The sample consisted of 2,509 individuals (40% female). Mean age of participants at Wave 1 was 47.6 years ($SD = 15.8$). The age range across waves was 14–89 years. Data on study variables were available for 2,509 individuals at Wave 1, 1,162 individuals at Wave 2, and 725 individuals at Wave 3. To investigate the potential impact of attrition, we compared individuals who did and did not participate in the most recent wave of data collection (2000) on study variables assessed at the first wave (1996). Participants who dropped out were more likely to be younger ($M_s = 46.6$ vs. 50.2 years, $d = -0.23$), have lower levels of education ($M_s = 2.28$ vs. 2.57, $d = -0.23$ [the scale ranged from 1 to 4; further information on the measures is provided later]), and have lower individual income ($M_s = 5.24$ vs. 5.48, $d = -0.10$; the scale ranged from 1 to 12) and household income ($M_s = 5.89$ vs. 6.13, $d = -0.12$; the scale ranged from 1 to 12); differences in self-esteem, relationship status, relationship satisfaction, employment status, job satisfaction, and health satisfaction were all nonsignificant. Given that all of these differences were small or nonsignificant, nonrepresentativeness due to attrition was not a serious concern in the present study. Overall, the sample was nationally representative on many, but not all, demographic variables; men and participants with higher education were overrepresented. An overview of gender and residence in West versus East Germany for the full sample and separately for six age groups is given in Table 1. A detailed description of the sample is available in German on the Internet (Schmal, Maes, & Schmitt, 1996).

Measures

Self-esteem. Self-esteem was assessed with a German translation of the 10-item Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1965), the most commonly used and well-validated measure of self-esteem (Robins et al., 2001), using a six-point scale ranging from 0 (strongly disagree) to 5 (strongly agree). The mean was 3.95 ($SD = 0.73$), averaged across waves. The alpha reliability was .84 at Wave 1 and .85 at Waves 2 and 3.

Relationship status. Relationship status was assessed with a dichotomous item: “Do you have a relationship partner?” At Waves 1–3, 74%, 81%, and 84% of the participants, respectively, were in a relationship.

Relationship satisfaction. Relationship satisfaction was assessed with five items (Schmitt, Maes, & Schmal, 1997). An example item is “I am [satisfied/dissatisfied] with the degree of tenderness and attention my partner shows to me.” Responses were measured on a six-point scale ranging from 0 (very dissatisfied) to 5 (very satisfied). The mean was 4.05 ($SD = 1.04$), averaged across waves. The alpha reliability was .93 at each wave.

Employment status. We used a dichotomous variable contrasting employed and unemployed participants. At Waves 1–3, 71%, 59%, and 61% of the participants, respectively, were employed.

Job satisfaction. Job satisfaction was assessed with five items (Schmitt et al., 1997). An example item is “I am [satisfied/dissatisfied] with the degree of success I have in my job.” Responses were measured on a six-point scale ranging from 0 (very dissatisfied) to 5 (very satisfied). The mean was 3.26 ($SD = 1.18$), averaged across waves. The alpha reliability was .81 at Wave 1, .86 at Wave 2, and .85 at Wave 3.

Individual income. Individual income was assessed with a 12-point measure ranging from 1 (less than 500 Deutsche Marks [U.S.$9,750] or more) to 12 (15,000 Deutsche Marks [U.S.$325] per month). The mean was 6.08 ($SD = 2.42$), averaged across waves.

Household income. Household income was assessed with a 12-point measure, ranging from 1 (less than 500 Deutsche Marks [U.S.$325] per month) to 12 (15,000 Deutsche Marks [U.S.$9,750] or more, per month). The mean was 5.48 ($SD = 1.28$), averaged across waves.

Health satisfaction. Health satisfaction was assessed with five items (Schmitt et al., 1997). An example item is “I am [satisfied/dissatisfied] with the state of my physical health.” Responses were measured on a six-point scale ranging from 0 (very dissatisfied) to 5 (very satisfied). The mean was 3.38 ($SD = 1.08$), averaged across waves. The alpha reliability was .89 at Wave 1, .91 at Wave 2, and .90 at Wave 3.

Education. We used a four-point measure of education (1 = degree after the obligatory school years [typically 9 or 10 years] or no degree, 2 = academic-track high school degree [typically 12 or 13 school years], 3 = degree from a “Fachhochschule” [a type of German college that focuses on applied sciences], and 4 = master’s degree from a university or a higher degree [such as a PhD, MD, etc.]). For most participants, because of their age, level of education was invariant across waves, so we used the highest degree that participants reported across waves as a time-invariant covariate in the analyses. The mean was 2.36 ($SD = 1.28$).

Statistical Analyses

Analyses were conducted using the Mplus 7.2 program (Muthén & Muthén, 2012). To deal with missing values, we used full information maximum likelihood estimation to fit models directly to the raw data, which produces less biased and more reliable results than do conven-
tional methods of dealing with missing data, such as listwise or pairwise deletion (Schafer & Graham, 2002; Widaman, 2006).

In the analyses of measurement models using item parcels as indicators, the estimator used was maximum likelihood. In measurement models using individual items as indicators, items were analyzed as ordered-categorical variables, using the mean- and variance-adjusted weighted least square estimator. Fit was assessed by the comparative fit index (CFI), the Tucker–Lewis index (TLI), and the root mean square error of approximation (RMSEA) on the basis of the recommendations of Hu and Bentler (1999) and MacCallum and Austin (2000). Hu and Bentler suggest that good fit is indicated by values greater than or equal to .95 for CFI and TLI and values less than or equal to .06 for RMSEA. To test for differences in model fit, we used the test of small difference in fit recommended by MacCallum, Browne, and Cai (2006, Program C). For these tests, statistical power was high with values above .99 (MacCallum et al., 2006, Program D).

In the analyses of growth curve models, which were based on individually varying times of measurement, the estimator used was robust maximum likelihood. For these models, CFI, TLI, and RMSEA were not available; therefore, fit was assessed using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). For AIC and BIC, absolute values cannot be interpreted, but when comparing models, lower values indicate better model fit (Burnham & Anderson, 2004).

### Results

#### Measurement Invariance of Self-Esteem

Before examining the life-span trajectory of self-esteem, we tested whether the self-esteem measure used showed measurement invariance. In testing for measurement invariance, we followed the procedures described by Widaman et al. (2010). First, we examined measurement invariance across time. An illustration of the measurement model used is provided in Figure 1. Item parcels were used as indicators of the latent factors because parcels produce more reliable latent variables than do individual items (Little, Cunningham, Shahar, & Widaman, 2002). Parcels were created in identical ways across waves, using the balancing technique recommended by Little, Rhemtulla, Gibson, and Schoemann (2013), and examined as continuous variables. The three latent self-esteem factors were allowed to correlate across waves. Moreover, the error variances of each parcel were allowed to correlate across waves to control for bias due to parcel-specific variance (Cole & Maxwell, 2003). In the measurement models, the mean and variance of the first latent factor were fixed to 0 and 1, respectively, so that all loadings and intercepts could be estimated. The first measurement model (i.e., Model 1) included configural invariance. The fit of this model was good (see Table 2). Models 2 and 3 tested weak and strong invariance by progressively constraining the loadings and intercepts of indicators, respectively, to be equal across waves. As indicated by the test of small difference in fit (MacCallum et al., 2006), the constraints did not significantly worsen model fit, suggesting strong measurement invariance across time.

Next, we examined measurement invariance across birth cohorts, using multigroup models. We created six cohorts using age at Wave 1 (i.e., 14–29, 30–39, 40–49, 50–59, 60–69, and 70 years and older; the youngest and oldest cohorts covered more than 10 years due to restrictions in sample size). Model 4 tested for configural invariance across cohorts, accounting for strong invariance across time (i.e., building on Model 3). The fit of this model was good (see Table 2). Models 5 and 6 tested weak and strong invariance across cohorts by progressively constraining the loadings and intercepts of indicators. As indicated by the test of small difference in fit, the constraints did not significantly worsen model fit, suggesting strong measurement invariance across birth cohorts.

Given that the use of item parcels is not unequivocally supported in the methodological literature (for the debate, see, e.g., Little et al., 2013; Marsh, Lüdtke, Nagengast, Morin, & Von Davier, 2013), we also tested for measurement invariance on the item level. We accounted for the Likert-type response scale of the RSE by using item factor analysis with categorical indicators (Wirth & Edwards, 2007). The model was similar to the parcel-level measurement model described earlier, except that the item-level model required including wave-specific method factors that accounted for bias due to positive and negative wording of the items (Marsh, Scalas, & Nagengast, 2010). Both the positive and negative wording factors were correlated across waves, but positive wording factors were uncorrelated with negative wording factors, and all wording factors were uncorrelated with the self-esteem factors. The fit of the models testing for measurement invariance across waves is shown in Table 2. Model 7 included configural invariance, and Models 8 and 9 tested weak and strong invariance by progressively constraining the loadings and thresholds of indicators, respectively, to be equal across waves. The fit of the models was good. Moreover, as indicated by the test of small difference in fit, the constraints did not significantly worsen model fit, suggesting strong measurement invariance across time. We also examined multigroup item-level models to test for measurement invariance across cohorts; however, none of these multigroup models converged, which was likely attributable to the computational complexity of item factor analysis. Nevertheless, given the findings for the full sample, the item-level analyses

---

**Figure 1.** Measurement model of self-esteem. At each wave, self-esteem was measured by three-item parcels (e.g., SE1a–SE1c at Wave 1). Error variances of parcels (denoted as e1–e9) were correlated across waves to control for bias due to parcel-specific variance (Cole & Maxwell, 2003).
In a similar manner to the measurement models described earlier, self-esteem was examined as a latent construct, as shown in Figure 1. The model accounted for strong measurement invariance. For scaling of the latent self-esteem scores, we followed the standardization method of Ferrer, Balluerka, and Widaman (2008). Because in the measurement models the mean and the variance of the first latent self-esteem factor were fixed to 0 and 1, respectively (resulting in a standardized metric in which self-esteem scores had a mean of 0 and a standard deviation of 1 relative to Wave 1), this standardized metric could be carried over to the growth models by constraining the first factor loading to the parameter estimate from the measurement model with strong invariance (Ferrer et al., 2008). Given that we examined the sample as a whole (but did not use a multigroup model with separate cohorts), we used the loading estimate from Model 3, which was 0.705. 

We estimated a model with an intercept only and linear, quadratic, and cubic growth curve models (Preacher et al., 2008). Because the slope loadings were based on age rather than the three measurement occasions, it was possible to estimate relatively complex trajectories. For the growth curve analyses, age was centered at 50 years and rescaled by the factor $10^{-2}$ to avoid numerically small estimates related to slope factors and to yield greater precision for these estimates. According to AIC, the quadratic model had the best fit to the data, whereas according to BIC, the linear model had the best fit (see Table 3). We therefore considered additional evidence for model selection. First, we examined the means and variances of the latent growth factors. In the linear model, one of the coefficients was nonsignificant—specifically, the variance of the linear slope ($p = .110$). Likewise, in the quadratic model, one of the coefficients was nonsignificant—specifically, the variance of the quadratic slope ($p = .084$). Although these coefficients did not meet an alpha level of .05, the means and variances of the growth factors overall suggested that both linear and quadratic slope factors were needed to capture the general trend and individual differences in the self-esteem trajectory. Second, we examined linear and quadratic models of the cross-sectional age differences in self-esteem at Wave 1; the results clearly favored the quadratic over the linear model.

Table 2

<table>
<thead>
<tr>
<th>Fit of Self-Esteem Measurement Models</th>
<th>$\chi^2$</th>
<th>df</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA [90% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Models using item parcels as indicators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invariance across waves</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Configural invariance</td>
<td>18.2</td>
<td>15</td>
<td>1.000</td>
<td>.999</td>
<td>.009 [.000, .022]</td>
</tr>
<tr>
<td>2. Weak invariance</td>
<td>19.9</td>
<td>19</td>
<td>1.000</td>
<td>1.000</td>
<td>.004 [.000, .018]</td>
</tr>
<tr>
<td>3. Strong invariance</td>
<td>41.5</td>
<td>23</td>
<td>.998</td>
<td>.996</td>
<td>.018 [.009, .027]</td>
</tr>
<tr>
<td>Invariance across cohorts*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Configural invariance</td>
<td>171.0</td>
<td>140</td>
<td>.996</td>
<td>.994</td>
<td>.023 [.006, .034]</td>
</tr>
<tr>
<td>5. Weak invariance</td>
<td>194.5</td>
<td>150</td>
<td>.994</td>
<td>.991</td>
<td>.027 [.014, .037]</td>
</tr>
<tr>
<td>6. Strong invariance</td>
<td>207.3</td>
<td>160</td>
<td>.994</td>
<td>.991</td>
<td>.027 [.015, .036]</td>
</tr>
<tr>
<td><strong>Models using items as indicators</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invariance across waves</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Configural invariance</td>
<td>692.8</td>
<td>336</td>
<td>.990</td>
<td>.987</td>
<td>.021 [.018, .023]</td>
</tr>
<tr>
<td>8. Weak invariance</td>
<td>667.8</td>
<td>354</td>
<td>.991</td>
<td>.989</td>
<td>.019 [.017, .021]</td>
</tr>
<tr>
<td>9. Strong invariance</td>
<td>929.3</td>
<td>452</td>
<td>.987</td>
<td>.987</td>
<td>.021 [.019, .022]</td>
</tr>
</tbody>
</table>

* Models 4–6 are multigroup models, accounting for strong invariance across waves.

Life-Span Trajectory of Self-Esteem

Next, we examined the life-span trajectory of self-esteem by using latent growth curve models that captured the development of self-esteem across the entire age range represented in the sample. Although each participant provided data for at most three age points, the complete life-span trajectory was constructed using the information from all participants simultaneously. This approach is based on the assumption (which is tested later) that a common trajectory can be modeled across all cohorts included in the sample (Duncan et al., 2006; Preacher et al., 2008). Given that the measurement was asynchronous across age (i.e., the data are organized by waves, but we were interested in another metric of time—specifically, the individual’s age at each wave), we used individual slope loadings, following the recommendations of Mehta and West (2000), Bollen and Curran (2006), and Preacher et al. (2008).

In a similar manner to the measurement models described earlier, self-esteem was examined as a latent construct, as shown in Figure 1. The model accounted for strong measurement invariance. For scaling of the latent self-esteem scores, we followed the standardization method of Ferrer, Balluerka, and Widaman (2008). Because in the measurement models the mean and the variance of the first latent self-esteem factor were fixed to 0 and 1, respectively (resulting in a standardized metric in which self-esteem scores had a mean of 0 and a standard deviation of 1 relative to Wave 1), this standardized metric could be carried over to the growth models by constraining the first factor loading to the parameter estimate from the measurement model with strong invariance (Ferrer et al., 2008). Given that we examined the sample as a whole (but did not use a multigroup model with separate cohorts), we used the loading estimate from Model 3, which was 0.705.

We estimated a model with an intercept only and linear, quadratic, and cubic growth curve models (Preacher et al., 2008). Because the slope loadings were based on age rather than the three measurement occasions, it was possible to estimate relatively complex trajectories. For the growth curve analyses, age was centered at 50 years and rescaled by the factor $10^{-2}$ to avoid numerically small estimates related to slope factors and to yield greater precision for these estimates. According to AIC, the quadratic model had the best fit to the data, whereas according to BIC, the linear model had the best fit (see Table 3). We therefore considered additional evidence for model selection. First, we examined the means and variances of the latent growth factors. In the linear model, one of the coefficients was nonsignificant—specifically, the variance of the linear slope ($p = .110$). Likewise, in the quadratic model, one of the coefficients was nonsignificant—specifically, the variance of the quadratic slope ($p = .084$). Although these coefficients did not meet an alpha level of .05, the means and variances of the growth factors overall suggested that both linear and quadratic slope factors were needed to capture the general trend and individual differences in the self-esteem trajectory. Second, we examined linear and quadratic models of the cross-sectional age differences in self-esteem at Wave 1; the results clearly favored the quadratic over the linear model.

Table 3

<table>
<thead>
<tr>
<th>Fit of Basic Growth Curve Models of Self-Esteem</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept only</td>
<td>25,017.7</td>
<td>25,175.0</td>
</tr>
<tr>
<td>Linear change</td>
<td>24,983.6</td>
<td>25,158.4</td>
</tr>
<tr>
<td>Quadratic change</td>
<td>24,972.7</td>
<td>25,170.9</td>
</tr>
<tr>
<td>Cubic change</td>
<td>24,982.3</td>
<td>25,209.6</td>
</tr>
</tbody>
</table>

* Note. For AIC (Akaike information criterion) and BIC (Bayesian information criterion), lower values indicate better model fit.
The results of the cross-sectional analyses are reported in detail in the supplemental materials. Third, as reviewed in the introduction, previous research suggests that a quadratic trajectory is a better model of the life-span trajectory of self-esteem than is a linear trajectory. Fourth, the methodological literature suggests that in situations in which the true model is not included in the set of models tested, AIC might be more appropriate than BIC to select the best-fitting model (Burnham & Anderson, 2004). Given that the models tested were only approximations of the true trajectory, in the present research, AIC might be the more useful criterion for model selection. For these reasons, we favored the quadratic model over the linear model.

The average predicted trajectory of self-esteem for the full sample is shown in Figure 2. Self-esteem tended to increase during adolescence, young adulthood, and middle adulthood—reaching a peak at age 60—and then to decline in old age. There was about a 1/5 standard deviation increase (\(d = 0.25\)) from age 14 to 60 and about a 1/5 standard deviation decrease (\(d = -0.22\)) from age 60 to 89 years. Overall, the trajectory was similar to the life-span trajectory found in previous longitudinal studies (Orth et al., 2010, 2012).

### Testing for Generational Changes in the Self-Esteem Trajectory

Next, we tested for cohort effects on the trajectory of self-esteem. As discussed in the introduction, it is possible that sociocultural changes during the 20th century have influenced the life-span trajectory of self-esteem. To disentangle the effects of aging and cohort, we estimated a conditional growth curve model in which the growth factors were regressed on birth cohort. The results indicated that birth cohort did not significantly predict any of the growth factors (including the intercept factor, linear slope, and quadratic slope; see Table 4). We concluded that there were no systematic generational changes in the level and shape of the self-esteem trajectory across the birth cohorts included in the present sample. Thus, the evidence suggests that modeling a single coherent trajectory across the observed age range—as described earlier—was appropriate. These findings are consistent with the results of previous cohort-sequential longitudinal studies (Erol & Orth, 2011; Orth et al., 2010, 2012).

### Table 4

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Intercept</th>
<th>Linear slope</th>
<th>Quadratic slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort</td>
<td>0.06 (0.05)</td>
<td>-0.27 (0.34)</td>
<td>0.02 (0.52)</td>
</tr>
<tr>
<td>Gender</td>
<td>-0.03 (0.05)</td>
<td>-0.19 (0.26)</td>
<td>-2.33 (1.35)</td>
</tr>
<tr>
<td>Education</td>
<td>0.05 (0.03)</td>
<td>-0.24 (0.13)</td>
<td>0.15 (0.86)</td>
</tr>
<tr>
<td>Residence</td>
<td>0.03 (0.06)</td>
<td>-0.28 (0.28)</td>
<td>-0.26 (1.71)</td>
</tr>
</tbody>
</table>

**Note.** The table reports unstandardized regression coefficients, with standard errors in parentheses (all coefficients were nonsignificant; standardized coefficients are not available). For the analyses, the age variable was rescaled by the factor \(10^{-2}\) to avoid numerically small estimates related to slope factors and to yield greater precision for these estimates. The original age scaling can be recovered by multiplying the regression coefficients of the linear slope by \(10^{-2}\) and the regression coefficients of the quadratic slope by \(10^{-4}\). The regression coefficients of the intercept were not affected by the rescaling of the age variable. Gender and residence were dummy coded, and cohort and education were centered for the analyses.

### Moderators of the Self-Esteem Trajectory

Then, we estimated conditional growth curve models to examine the moderating effects of gender, education, and residence (West vs. East Germany) on the self-esteem trajectory. Gender and residence were dummy coded and education centered for these analyses. For all variables, none of the effects on the intercept or on the linear and quadratic slopes were significant (see Table 4). We also tested for two-way and three-way interaction effects between these covariates; all interaction effects were nonsignificant.

Finally, we examined the effects of time-varying covariates (TVCs) on the self-esteem trajectory. Given that variables such as relationship status and income may change over time, it is important that the model accounts for potential changes in the TVCs. The generic model that was used for the analyses is shown in Figure 3. In this model, self-esteem at specific waves is explained simultaneously by the growth factors and a repeatedly measured TVC (Bollen & Curran, 2006; Preacher et al., 2008). The model estimates the effect of the TVC on self-esteem while simultaneously controlling for systematic growth in self-esteem; in other words, the model tests whether the TVC predicts deviations from the expected self-esteem trajectory. Dichotomous TVCs were dummy coded and continuous TVCs converted to \(z\) scores for the analyses. The results indicated that relationship satisfaction, employment status, job satisfaction, household income, and health satisfaction had significant TVC effects (see Table 5). Given that self-esteem and continuous TVCs were measured in a standardized metric and dichotomous TVCs were dummy coded, the unstandardized coefficients reported in Table 5 can be readily interpreted. For example, for relationship satisfaction the coefficient was 0.19, indicating that when relationship satisfaction of two individuals differs by 1 standard deviation, a difference of 0.19 standard deviations in self-esteem is predicted. Another example is employment status; its coefficient was 0.15, indicating that at each wave, the self-esteem of employed individuals was predicted to be 0.15 standard deviations higher than the self-esteem of unemployed individuals.

![Figure 2](image-url)  
*Figure 2.* Average predicted trajectory of self-esteem from ages 14 to 89 years.
Discussion

In this study, we examined the development of self-esteem across the life span using longitudinal data from a large sample. Latent growth curve analyses indicated that self-esteem follows a quadratic trajectory across the life span, increasing during adolescence, young adulthood, and middle adulthood—reaching a peak at age 60 years—and then declining in old age. No cohort effects on average levels of self-esteem or on the shape of the trajectory were found. Moreover, the trajectory did not differ across gender and level of education. However, the results suggested that employment status, household income, and satisfaction in the domains of work, relationships, and health contribute to a more positive life-span trajectory of self-esteem.

Implications of the Findings

The present research replicates in a German sample the curvilinear trajectory of self-esteem across the life span found in two previous studies using samples from the United States (Orth et al., 2010, 2012). Overall, the shape of the trajectory was similar in all three studies. Across studies, the peak of the trajectory was estimated to occur between ages 51 and 60 years. However, effect sizes of the increases and decreases over the life course should be considered to compare the findings across studies more closely. Whereas the increase from adolescence (or young adulthood [the sample examined in Orth et al., 2010, did not include participants younger than 25 years of age]) to midlife corresponded to about 1/3 to ½ of a standard deviation in all studies, the effect sizes of the decrease from midlife to old age differed more strongly. Whereas in previous studies the decline corresponded to about 2/3 of a standard deviation (Orth et al., 2010, 2012), in the present study there was a smaller decline of about 1/5 of a standard deviation. A possible explanation is that the sample in the present study came from a different country (Germany) than the samples of previous studies (United States). Given that negative changes in health and socioeconomic status partially explain the self-esteem drop in old age (Orth et al., 2010), cross-cultural differences in covariates might account for diverging self-esteem trajectories among older persons.
adults. A promising way to gain more precise knowledge about the magnitude of self-esteem change in old age, and about significant moderators of change, is provided by meta-analytic methods. Although in a recent meta-analysis few samples from old age were available (i.e., there were four samples of adults 60 years or older; Huang, 2010), it is likely that the number of studies will increase in the coming years (e.g., Wagner et al., 2014). Nevertheless, with regard to the overall pattern of the life-span trajectory, the available data provide converging evidence that self-esteem increases from adolescence to midlife and then decreases in old age.

The findings suggest that predictors from several life domains are important in explaining interindividual variability in the self-esteem trajectory. The TVC analyses show that satisfaction in the relationship, work, and health domains has significant effects on self-esteem, over and above normative age trends in self-esteem. Moreover, in addition to these subjective indicators of well-being, significant effects emerged for two objective measures of life outcomes (i.e., employment status and household income), which showed the same pattern as the subjective measures. With regard to gender differences, the present findings converge with findings from previous studies. In this study, men and women did not differ in the level and shape of the trajectory; similarly, in previous longitudinal studies, the gender difference was nonsignificant (Erol & Orth, 2011; Orth et al., 2012) or small (Orth et al., 2010; Wagner, Lüdtke et al., 2013).

Further evidence of the robustness of the findings is provided by the fact that the trajectory replicated across the two subsamples of participants who had lived continuously in either East or West Germany. This finding is important, because these two parts of Germany were separate states from 1949 to 1990, with a low level of exchange between the populations (in terms of migration and travel) and important political, economic, and social differences, so two relatively distinct cultural systems emerged (cf. Schmitt & Maes, 1998, 2002). Despite these differences, however, East and West Germans showed the same pattern of self-esteem development, providing further evidence for the cross-cultural generalizability of the curvilinear pattern of the life-span trajectory.

The findings of the present study are consistent with hypotheses about the life-span development of self-esteem derived from neo-socioanalytic theory (Roberts & Wood, 2006; see also Roberts et al., 2008). As described in the introduction, the theory proposes that individual characteristics such as self-esteem can change across the whole life course. Specifically, the theory suggests that self-esteem shows systematic increase during young and middle adulthood, because people typically invest strongly in important social roles during these developmental periods, which—according to the social investment principle and the maturity principle—strengthens adaptive personality characteristics, including self-esteem. For old age, the theory suggests that self-esteem might no longer show increase but, rather, decrease, as old age frequently includes negative changes in social roles. The present findings correspond to the depicted shape of the life-span trajectory and suggest that the degree of success in social roles in the relationship and work domains moderates the self-esteem trajectory. In future research, it would be highly interesting to test in more detail how social roles influence the individual’s self-esteem by assuming specific roles (such as the roles of spouse, parent, or supervisor), committing oneself to a role, investing in a role, or dropping out of important roles.

The present study provides evidence with regard to the debate about whether more recent generations experience a higher level or a steeper increase in self-esteem than did previous generations (Trzesniewski & Donnellan, 2010; Twenge & Campbell, 2001). Although the hypothesis that sociocultural changes (such as greater cultural importance placed on self-esteem and grade inflation in education) have led to generational increases in self-esteem is plausible, the present study suggests that the typical life-span trajectory has not changed across the generations born in the past century in Germany (neither the overall level nor the shape of the trajectory), which is consistent with the findings from previous longitudinal studies on the life-span trajectory of self-esteem (Orth et al., 2010, 2012).

Limitations and Future Directions

One limitation of the present study is that the data were collected across a 4-year interval only. However, although a longer study period would clearly have been desirable, the longitudinal design of the data is valuable because it allowed examination of patterns of intraindividual change, which is essential for the study of developmental trajectories. Moreover, given that the trajectories were not confounded by cohort effects, the present data provide valid insights into the trajectory that would be found if individuals were followed across the full time span from adolescence to old age. Another limitation is the relatively low response rate, which might have introduced sample selectivity. Nevertheless, the sample was nationally representative with regard to many demographic characteristics (Schmal et al., 1996). Moreover, the mean self-esteem score of the present sample was very close to normative data for the German version of the RSE, on the basis of a nationally representative sample of 4,988 individuals ages 14 to 92 years (Roth, Decker, Herzberg, & Brähler, 2008). A further limitation is that parts of the sample dropped out of the study between the first and last assessment. However, the fact that attrition effects on all study variables were small or nonsignificant (in particular, there was no attrition effect for self-esteem) suggests that the conclusions of the present research are not biased by attrition. In addition, a disadvantage of the study is that the sample did not include participants younger than 14 years of age; thus, the present study provides no information on the self-esteem trajectory in childhood. Although theory and empirical studies suggest that self-esteem declines during middle childhood (e.g., Robins et al., 2002), the evidence from longitudinal studies is inconsistent (Cole et al., 2001; Huang, 2010; Kuzucu, Bontempo, Hofer, Stallings, & Piccinin, 2014). Future research should therefore focus more closely on the self-esteem trajectory during childhood.

Another limitation is that in the present research only linear and curvilinear models of self-esteem development were tested. Although the present results overall support the quadratic model, and although the observed means (as reported in the supplemental materials) suggested that a quadratic function provides a useful approximation of age differences in self-esteem, it is likely that the true change function is more complex. For example, it is possible that during late young adulthood and early middle adulthood (i.e., from about ages 30 to 50 years), the increase in self-esteem is smaller than predicted by the quadratic model (see the observed means in the supplemental materials). Thus, in future research, it would be desirable to examine more complex change functions.
increase statistical power when testing for complex functions, future studies should include a larger number of repeated assessments than included in the present study. Nevertheless, despite the fact that the quadratic curve is only an approximation of self-esteem development from adolescence to old age, the present research advances the field by providing evidence on the general life-span trajectory that converges with evidence from the previous longitudinal studies on self-esteem across the life span (Orth et al., 2010, 2012).

It is worth noting that the present research used an explicit measure of self-esteem. Although, at present, explicit measures are considered the most valid measures of self-esteem (Buhrmester, Blanton, & Swann, 2011), in future research, it would be interesting to study self-esteem development and cohort effects on self-esteem using implicit measures of self-esteem (e.g., Greenwald & Farnham, 2000). Whether studies of implicit self-esteem would yield the same pattern of results as the present research is currently unknown. In addition, in future research, it would be interesting to distinguish between secure and defensive self-esteem (a distinction that is not possible on the basis of the self-esteem measure used in the present study). Given the differing reactions of individuals with secure versus defensive self-esteem (Baumeister, Smart, & Boden, 1996; Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003), it is important to know whether these forms of self-esteem show the same or different normative developmental trajectories. Some evidence is provided by a cross-sectional study that examined age differences in the level, contingency, and intraindividual stability of self-esteem (Meier, Orth, Denissen, & Kühnle, 2011). The findings suggested that during young and middle adulthood, self-esteem tends to become not only higher but also less contingent and more stable (i.e., more secure).

Strengths of the present research include the large sample size, the availability of a well-validated measure of self-esteem (i.e., the full 10-item RSE), and the presence of strong measurement invariance of the self-esteem measure across assessments and birth cohorts. Moreover, although the sample was not a strict national probability sample, the study included random samples from geographically diverse regions of Germany, which strengthens the generalizability of the results. Also, in this study, self-esteem was examined as a latent construct, which allowed for controlling for the confounding influence of measurement error, increasing the validity of the findings.

In conclusion, the present research improves our understanding of self-esteem development by providing longitudinal evidence on the shape of the life-span trajectory; by showing that the curvilinear trajectory previously found in U.S. samples replicates in the present German sample and, moreover, among individuals who had been living continuously in the two former states of West and East Germany; by providing evidence that the average level and shape of the self-esteem trajectory did not change across cohorts born in Germany between about 1910 to 1980; and by showing that level of success and well-being in the work, relationship, and health domains are important predictors of individual differences in the self-esteem trajectory. The findings have significant implications, because they call attention to developmental stages in which individuals may be vulnerable as a result of low self-esteem (such as adolescence and old age) and to factors that predict successful versus problematic developmental trajectories.

References


This document is copyrighted by the American Psychological Association or one of its allied publishers. This article is intended solely for the personal use of the individual user and is not to be disseminated broadly.