Development of Self-Esteem From Age 4 to 94 Years: A Meta-Analysis of Longitudinal Studies

Ulrich Orth, Ruth Yasemin Erol, and Eva C. Luciano

University of Bern

Orth, U., Erol, R. Y., & Luciano, E. C. (2018). Development of self-esteem from age 4 to 94 years: A meta-analysis of longitudinal studies. *Psychological Bulletin, 144*, 1045-1080. http://dx.doi.org/10.1037/bul0000161

Author Note

Ulrich Orth, Ruth Yasemin Erol, and Eva C. Luciano, Department of Psychology, University of Bern.

Correspondence concerning this article should be addressed to Ulrich Orth, Department of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland. E-mail: ulrich.orth@psy.unibe.ch

Abstract

To investigate the normative trajectory of self-esteem across the life span, this meta-analysis synthesizes the available longitudinal data on mean-level change in self-esteem. The analyses were based on 331 independent samples, including data from 164,868 participants. As effect size measure, we used the standardized mean change *d* per year. The mean age associated with the effect sizes ranged from 4 to 94 years. Results showed that average levels of self-esteem increased from age 4 to 11 years (cumulative $d = 0.34$; cumulative *d*s are relative to age 4), remained stable from age 11 to 15, increased strongly until age 30 (cumulative $d = 1.05$), continued to increase until age 60 (cumulative $d = 1.30$), peaked at age 60 and remained constant until age 70, declined slightly until age 90 (cumulative $d = 1.15$), and declined more strongly until age 94 (cumulative $d = 0.76$). Moderator analyses were conducted for the full set of samples and for the subset of samples between ages 10 to 20 years. Although the measure of self-esteem accounted for differences in effect sizes, the moderator analyses suggested that the pattern of mean-level change held across gender, country, ethnicity, sample type, and birth cohort. The meta-analytic findings clarify previously unresolved issues about the nature and magnitude of self-esteem change in specific developmental periods (i.e., childhood, adolescence, and old age) and draw a much more precise picture of the life-span trajectory of self-esteem.

Keywords: self-esteem, life-span development, mean-level change, longitudinal studies, meta-analysis

Public Significance Statement

This meta-analysis shows that people’s self-esteem changes in systematic ways over the life course. On average, self-esteem increases in early and middle childhood, remains constant (but does not decline) in adolescence, increases strongly in young adulthood, continues to increase in middle adulthood, peaks between age 60 and 70 years, and then declines in old age, with a sharper drop in very old age. The pattern of findings holds across gender, country, ethnicity, and birth cohort.

Development of Self-Esteem From Age 4 to 94 Years:

A Meta-Analysis of Longitudinal Studies

Self-esteem is by no means an immutable characteristic of individuals. People experience changes in their self-esteem, both in terms of temporary boosts or drops in their feelings of self-worth and in terms of long-term increases or declines in their general level of self-esteem. For example, successes at school, work conflicts, or harmonious family events may cause transient fluctuations in self-esteem (J. Crocker & Luhtanen, 2003; J. Crocker & Wolfe, 2001). Also, stressful life events, such as a criminal victimization, and life transitions, such as beginning a satisfying romantic relationship, may lead to sustained changes in self-esteem (Luciano & Orth, 2017; Orth & Luciano, 2015). But does self-esteem follow a typical, normative pattern of change across the human life course?

For a long time, the literature suggested that self-esteem does not show systematic change at any age (Wylie, 1979). However, over the past one or two decades research has challenged this notion (for an early cross-sectional study, see Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002). Longitudinal studies have generally suggested that, on average, self-esteem increases from adolescence to middle adulthood, peaks at about age 50 to 60 years, and then decreases in old age (for reviews, see Orth & Robins, 2014, in press; Trzesniewski, Donnellan, & Robins, 2013). Thus, a growing body of evidence suggests that self-esteem follows a normative trajectory across the life span, as has been found for many other personality characteristics such as the Big Five personality traits (Lucas & Donnellan, 2011; Roberts, Walton, & Viechtbauer, 2006; Soto, John, Gosling, & Potter, 2011) and people’s general sense of control (Specht, Egloff, & Schmukle, 2013).

However, important questions about the pattern of self-esteem development are still unanswered because the evidence has been inconsistent for some age groups, and also because some of these issues can hardly be resolved by any single study but can better be addressed with meta-analytic methods (below, we discuss the unresolved issues in detail). In the present research, we therefore synthesize the available data on mean-level change in self-esteem across the life course, with the goal of drawing a precise picture of the normative self-esteem trajectory from childhood to old age. Moreover, we test for possible moderators of mean-level change in self-esteem. Mean-level change is defined as change in the average level of a construct between two repeated assessments (e.g., separated by one year) of the same sample. When mean-level change is mapped on age, it is also referred to as normative change.

Understanding the life-span development of self-esteem is important because research suggests that self-esteem truly matters for people's lives. Although researchers have debated whether self-esteem has any influence on important life outcomes (Baumeister, Campbell, Krueger, & Vohs, 2003; Krueger, Vohs, & Baumeister, 2008; Swann, Chang-Schneider, & McClarty, 2007, 2008), a growing body of evidence supports the notion that self-esteem does have consequences for people's lives (for reviews, see Donnellan, Trzesniewski, & Robins, 2011; Orth, 2017b; Orth & Robins, 2014). In particular, prospective studies suggest that self-esteem influences people's success and well-being in the domains of social relationships (M. D. Johnson & Galambos, 2014; Marshall, Parker, Ciarrochi, & Heaven, 2014; Mund, Finn, Hagemeyer, Zimmermann, & Neyer, 2015; Orth, Robins, & Widaman, 2012), school and education (Trzesniewski et al., 2006; von Soest, Wichstrom, & Kvalem, 2016), work (Kuster, Orth, & Meier, 2013; Orth et al., 2012; Trzesniewski et al., 2006; von Soest et al., 2016), physical health (Orth et al., 2012; Trzesniewski et al., 2006), and mental health (Orth, Robins, Meier, & Conger, 2016; Orth, Robins, & Roberts, 2008; Sowislo & Orth, 2013; Sowislo, Orth, & Meier, 2014; Steiger, Allemand, Robins, & Fend, 2014; Wouters et al., 2013; for a review, see Orth & Robins, 2013). Importantly, research in this field allows for relatively strong conclusions because many of the studies used large and representative samples, controlled for prior levels of the outcomes, and controlled for confounding factors such as gender, socioeconomic status, intelligence, and life events.

Before reviewing the theoretical and empirical background on mean-level change in self-esteem, we provide a definition of the construct. Self-esteem is defined as a person's subjective evaluation of his or her worth as a person (e.g., Donnellan et al., 2011; MacDonald & Leary, 2012). Thus, self-esteem is by definition a subjective construct and does not necessarily reflect objective characteristics of the person, or how the person is seen by others. It is therefore important to distinguish self-esteem from narcissism, as both constructs involve positive self-evaluations (Brummelman, Thomaes, & Sedikides, 2016; Orth & Luciano, 2015). Self-esteem includes feelings of self-acceptance and a positive attitude toward the self, but does not necessarily imply that the individual feels superior to others (Rosenberg, 1965). In contrast, narcissism is characterized by feelings of grandiosity and superiority, self-centeredness, sense of entitlement, willingness to exploit others, and lack of empathy (Ackerman et al., 2011; Bosson et al., 2008; Morf & Rhodewalt, 2001). Thus, whereas narcissism is related to antisocial behavior and a negative view of others, high self-esteem is compatible with a prosocial, positive attitude towards others (Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005; Paulhus, Robins, Trzesniewski, & Tracy, 2004; Tracy, Cheng, Robins, & Trzesniewski, 2009).

Theoretical Perspectives on Self-Esteem Development

To date, no theories that focus on the life-span development of self-esteem have been proposed. In the following, we therefore review theoretical perspectives that address specific developmental periods (e.g., childhood and adolescence) and related constructs (e.g., broader personality traits).

Susan Harter's work (e.g., Harter, 2006a, 2006b, 2006c, 2012a) focuses on the development of self-esteem during childhood and adolescence. With regard to mean-level change, Harter identified several processes that might cause a decline in self-esteem from early childhood to middle childhood (i.e., from about age 4 to 8 years). Ironically, although children show important social-cognitive advances during this period, the very advances may compromise children's self-esteem or, as Harter has put it, represent "liabilities" (Harter, 2006b, 2012a). For example, four-year old children are not yet able to distinguish between their actual and their ideal competences. When interviewed about their self-views, children of this age describe their ideal self rather than their actual self (Harter, 2006b). Later, when children learn to discriminate between their actual and ideal characteristics, this process leads to reduced positivity of self-descriptions. Young children also lack the ability to use social comparison information in their self-evaluation (Ruble, Boggiano, Feldman, & Loebl, 1980). Again, when children overcome this cognitive limitation in middle childhood, it might lead to a decline in self-esteem (Harter, 2006a). A third social-cognitive process is perspective taking. Whereas young children are not yet able to infer how the self is evaluated by significant others (e.g., parents and caregivers), by about age 8 children have considerably improved their social perspective-taking skills (Harter, 2006a). Thus, beginning in middle childhood, children's self-views are more strongly influenced by how others perceive them. Because not all persons with whom children regularly interact see them positively, many children will experience somewhat lower levels of self-esteem (Harter, 2006a). According to Harter (2006c), self-esteem further declines at the transition to adolescence, a decline that may be accounted for by even stronger emphasis on social comparison, less personal attention by teachers, and pubertal changes. In middle adolescence, however, self-esteem begins to recover, a process that may be explained by increases in personal autonomy, sense of control, and greater possibilities in choosing friends, peer groups, and activities that match the individual's personality (Harter, 2006c).

With regard to adulthood, important background for understanding the development of self-esteem is provided by theory on personality development (for a review of theories in this field, see Specht et al., 2014). In particular, neo-socioanalytic theory allows deriving hypotheses about mean-level change in self-esteem (Roberts & Wood, 2006; see also Roberts, Wood, & Caspi, 2008). This theory suggests that adults typically develop in the direction of more mature personality traits, especially during young adulthood. The reason is that individuals assume many social roles (e.g., the roles of relationship partner, parent, employee, etc.) and that social roles involve social expectations about role-congruent behavior. For most social roles, these expectations include conscientiousness, agreeableness, and emotional stability (i.e., mature personality characteristics). Given that most individuals are committed to satisfying these expectations, many individuals gradually improve on these traits, a mechanism that has been called the social investment principle (Roberts & Wood, 2006; Roberts et al., 2008). Especially during young adulthood, individuals make the transition into many of these roles, by entering into working life, committing to a stable romantic relationship, having a first baby, and assuming additional social roles in the community (Hutteman, Hennecke, Orth, Reitz, & Specht, 2014). Given that mature personality traits are associated with higher self-esteem (Robins, Hendin, & Trzesniewski, 2001; Watson, Suls, & Haig, 2002) and given that self-esteem is itself a personality characteristic that improves functioning in social roles (Orth et al., 2012), neo-socioanalytic theory suggests that young adults show a relatively strong normative increase in self-esteem. During middle adulthood, most people further invest in their social roles, for example by taking on managerial roles at work, maintaining a satisfying relationship with their spouse or partner, and by helping their children to become responsible and independent adults (Hutteman et al., 2014). Nevertheless, research on personality development suggests that the developmental pace in the direction of maturity fades (Lucas & Donnellan, 2011; Roberts et al., 2006; Roberts et al., 2008; Terracciano, McCrae, Brant, & Costa, 2005), suggesting that the normative increase in self-esteem becomes smaller during middle adulthood.

For old age, however, neo-socioanalytic theory suggests that self-esteem no longer increases, but instead decreases, because old age frequently involves loss of social roles due to retirement and, possibly, widowhood. Correspondingly, research on the Big Five personality traits has shown negative changes in some traits in old age, such as conscientiousness (Lucas & Donnellan, 2011; Marsh, Nagengast, & Morin, 2013; Specht, Egloff, & Schmukle, 2011). In addition to changes in social roles, aging often leads to negative changes in other possible sources of self-esteem, such as social relationships, socioeconomic status, cognitive abilities, and health (Baltes & Mayer, 1999; Wagner, Gerstorf, Hoppmann, & Luszcz, 2013).

Research on subjective well-being may provide further relevant background for understanding the life-span development of self-esteem (Diener, Suh, Lucas, & Smith, 1999; Luhmann, 2017). For example, life satisfaction is stable, or tends to increase, during young and middle adulthood, but tends to decrease in old age (Baird, Lucas, & Donnellan, 2010; Gerstorf, Ram, Estabrook, et al., 2008; Gerstorf, Ram, Röcke, Lindenberger, & Smith, 2008; Mroczek & Spiro, 2005). With regard to positive affectivity, research generally suggests that it tends to increase, or remains stable, during adulthood, whereas negative affectivity decreases across large parts of adulthood and remains at a low level in old age (Carstensen, Pasupathi, Mayr, & Nesselroade, 2000; Charles, Reynolds, & Gatz, 2001; Helson & Soto, 2005; Mroczek & Kolarz, 1998). Finally, depression, which is a key indicator of negative affect, decreases from young adulthood to midlife, but tends to increase in old age (Kasen, Cohen, Chen, & Castille, 2003; Kessler, Foster, Webster, & House, 1992; Mirowsky & Kim, 2007; J. Wallace & O'Hara, 1992). Overall, however, it should be noted that the available evidence on mean-level change in well-being across the life span is relatively inconsistent (Luhmann, 2017).

From a theoretical perspective, it is useful to note that other fields of research suggest that self-esteem can show systematic, gradual change over the life span. First, behavioral genetic research indicates that both genes and environment account for variance in self-esteem, but that the influence of environmental factors is slightly greater than the influence of genetic factors (Bleidorn, Hufer, Kandler, Hopwood, & Riemann, 2018; McGuire, Neiderhiser, Reiss, Hetherington, & Plomin, 1994; Neiss, Sedikides, & Stevenson, 2002; Neiss et al., 2005). Thus, if environmental factors change as a function of age, then these factors could cause normative mean-level change in self-esteem. Second, although research suggests that the rank-order stability of self-esteem increases with age, peaking in young or middle adulthood (Trzesniewski, Donnellan, & Robins, 2003), estimates of the long-term stability of individual differences are far from unity, clearly allowing for the possibility of self-esteem change over time (Anusic & Schimmack, 2016; Donnellan, Kenny, Trzesniewski, Lucas, & Conger, 2012; Kuster & Orth, 2013; Wagner, Lüdtke, & Trautwein, 2016). Thus, even if self-esteem is relatively consistent across time and should be considered a personality trait, self-esteem change is possible.

Mean-Level Change in Self-Esteem Across the Life Span

Empirical Findings

As noted above, cross-sectional studies were the first to indicate that self-esteem changes in systematic ways across the life span (McMullin & Cairney, 2004; Pullmann, Allik, & Realo, 2009; Robins et al., 2002). For example, the pattern of age differences found by Robins et al. (2002) suggested that self-esteem decreases from childhood to adolescence, increases during young and middle adulthood, reaches a peak at age 65 years, and decreases into old age. However, given the cross-sectional nature of the data, these studies did not allow disentanglement of true developmental effects and cohort differences (Baltes, Cornelius, & Nesselroade, 1979). Thus, any observed pattern of cross-sectional age differences could be severely biased and convey a misleading picture of the developmental trajectory. In Robins et al. (2002), for example, participants who were in their 80s at the time of the study and who reported the lowest self-esteem scores of this sample may have had low self-esteem for their whole life because of, e.g., adverse parenting and economic conditions when they were young.

To address these interpretational problems of cross-sectional studies, longitudinal data are needed. As yet, three longitudinal studies have modeled the life-span trajectory of self-esteem (Orth, Maes, & Schmitt, 2015; Orth et al., 2012; Orth, Trzesniewski, & Robins, 2010). Importantly, all three studies used cohort-sequential designs, which allow separation of intraindividual change from cohort effects (Baltes et al., 1979). The studies included large samples with broad age ranges (i.e., at the first wave of assessment, age ranged at least from young adulthood to old age) and participants had been assessed multiple times across up to 16 years. Using latent growth modeling, the studies tested for competing models of the life-span trajectory, such as linear, quadratic, and cubic change (Duncan, Duncan, & Strycker, 2006; Preacher, Wichman, MacCallum, & Briggs, 2008). Taken together, the three studies suggested that self-esteem follows an inverted U-shape over the life span, with self-esteem increasing from adolescence to middle adulthood, peaking at about 50 to 60 years of age, and decreasing during old age. The increase from adolescence to middle adulthood corresponded to an effect size of about d = 0.30 to 0.50 (expressed as standardized mean change; Cohen, 1992). The decrease from middle adulthood to old age varied much more strongly across studies, with the effect size ranging from about d = −0.20 to −0.70. Overall, this pattern of the life-span trajectory is supported by longitudinal research on mean-level change during specific developmental periods, such as adolescence and young adulthood (Birkeland, Melkevik, Holsen, & Wold, 2012; Chung, Hutteman, van Aken, & Denissen, 2017; Chung et al., 2014; Erol & Orth, 2011; Galambos, Barker, & Krahn, 2006; Harris, Wetzel, Robins, Donnellan, & Trzesniewski, 2018; Kiviruusu, Huurre, Aro, Marttunen, & Haukkala, 2015; Wagner, Lüdtke, Jonkmann, & Trautwein, 2013; Zeiders, Umaña-Taylor, & Derlan, 2013) and old age (von Soest, Wagner, Hansen, & Gerstorf, 2017; Wagner, Gerstorf, et al., 2013; Wagner, Hoppmann, Ram, & Gerstorf, 2015; Wagner, Lang, Neyer, & Wagner, 2014).

Unresolved Issues

Although the studies cited above provide a relatively consistent picture of the general pattern of self-esteem change across large parts of the life span, the evidence is unclear for middle childhood, early adolescence, and old age.

As reviewed above, theory suggests that children experience decreases in their self-esteem when they transition from early to middle childhood (i.e., from about age 4 to 8 years), due to cognitive advances such as in the ability to use social comparison information and in perspective-taking abilities (Harter, 2006c; Ruble et al., 1980; Stipek & Tannatt, 1984). Moreover, the literature suggests that preschool children often have inflated self-views, rating themselves much more positively than they rate other children, but that this bias declines with age (Robins et al., 2002; Stipek & Tannatt, 1984). Some empirical studies support the notion that children's self-esteem declines in this age period (J. Eccles, Wigfield, Harold, & Blumenfeld, 1993; Marsh, 1989; Marsh, Barnes, Cairns, & Tidman, 1984; Stipek & Tannatt, 1984). However, because most studies in this developmental period used measures of domain-specific self-evaluations (e.g., evaluations of one's academic abilities, social competence, and physical appearance) but did not assess global self-esteem, more evidence is needed regarding the nature of the self-esteem trajectory in middle childhood.

Also, reviews of the literature suggest that self-esteem declines at the transition from childhood to early adolescence (i.e., at about age 11 to 13 years), with early adolescents experiencing a low point in self-esteem (Harter, 2006c; Robins et al., 2002). One possible cause of the self-esteem decline at this age is the educational transition from elementary school to middle school (although the precise timing, and even existence, of this transition varies across countries), with disruptions in the child's social network and even greater emphasis on social comparison and competition than before (Blyth, Simmons, & Carlton-Ford, 1983; J. S. Eccles et al., 1989; Harter, 2006c; Simmons, Blyth, Van Cleave, & Bush, 1979; Wigfield, Eccles, Mac Iver, Reuman, & Midgley, 1991). Another possible cause of the decline are the many physical, emotional, and social changes associated with puberty (Simmons et al., 1979; Simmons & Rosenberg, 1975). However, some empirical studies did not find evidence of decreasing self-esteem in this age group (Cole et al., 2001; Huang, 2010; Kuzucu, Bontempo, Hofer, Stallings, & Piccinin, 2014). Thus, it is unclear whether self-esteem decreases in early adolescence and reaches a low point, or whether it remains stable or even increases.

Old age is a third developmental period for which the evidence is unclear. Although many studies suggest that self-esteem declines after about age 65–70 years, the magnitude of change (i.e., the slope of the trajectory) differs strongly across studies. Whereas some studies found relatively large effect sizes in old age (Orth et al., 2012; Orth et al., 2010; Shaw, Liang, & Krause, 2010; von Soest et al., 2017), other studies indicated that self-esteem is stable in this age group (Wagner et al., 2014) or that there is at most a benign decrease (Coleman, Ivani-Chalian, & Robinson, 1993; Orth et al., 2015; Wagner, Gerstorf, et al., 2013; Wagner, Hoppmann, et al., 2015). Thus, meta-analytic synthesis may help to more reliably test whether self-esteem declines in old age and, if there is a decline, to estimate the degree of the decline and the average age at which the decline begins.

Moreover, even if the general uptrend of self-esteem from adolescence to middle adulthood has been replicated several times in longitudinal studies, the present meta-analysis will provide a more reliable picture of the precise shape of the trajectory across these developmental periods. An important strength of the meta-analytic method is the high statistical power and the estimation of effect sizes across a large number of samples.

Moderators of Mean-Level Change in Self-Esteem

Empirical Findings

The research on mean-level change in self-esteem reviewed above has shown that people differ significantly in the individual trajectory they follow. Evidence for individual differences is provided, e.g., by the variances of growth factors in studies using latent growth modeling (Orth et al., 2010; Wagner, Gerstorf, et al., 2013; Young & Mroczek, 2003) as well as by research employing growth mixture modeling, suggesting that there may be distinct trajectories (Birkeland et al., 2012; Morin, Maiano, Marsh, Nagengast, & Janosz, 2013; Mund & Neyer, 2016). The variability in patterns of self-esteem development raises the question as to which factors explain the individual trajectories.

Many longitudinal studies have tested for gender differences in mean-level change of self-esteem. Some studies found that there is a small gender difference in favor of boys and men (Orth et al., 2010; von Soest et al., 2016; Wagner, Lüdtke, et al., 2013), whereas in other studies the gender difference was nonsignificant (Erol & Orth, 2011; Orth et al., 2015; Orth et al., 2012). Overall, these findings correspond to meta-analyses of cross-sectional data, which suggested that men and women differ only slightly in their average levels of self-esteem (Kling, Hyde, Showers, & Buswell, 1999; Major, Barr, Zubek, & Babey, 1999; Zuckerman, Li, & Hall, 2016). However, even if there is a small gender difference in the self-esteem trajectory, research suggests that gender is not a direct cause but that associated factors, such as differential treatment of men and women in educational and work contexts, account for the effect (Zuckerman et al., 2016).

Few studies tested for ethnic differences in the self-esteem trajectory, based on samples from the United States (Erol & Orth, 2011; Orth et al., 2010; Shaw et al., 2010). Overall, the evidence suggests that ethnicity might moderate the life-span trajectory. When compared to Americans of European descent, African Americans reported a stronger increase in adolescence and young adulthood, but also a stronger decrease in old age (Orth et al., 2010; Shaw et al., 2010). Similarly, when compared to European Americans, Hispanic Americans experienced a stronger increase in adolescence and young adulthood, although starting at a lower level in adolescence (Erol & Orth, 2011). However, effect sizes of ethnic differences were typically small. Moreover, although in a large cross-sectional study ethnic groups (including White, Black, Hispanic, and Asians) differed in the level of self-esteem, the study suggested that the shape of the life-span trajectory replicated across ethnic groups (Robins et al., 2002).

Research studies have examined additional factors that may moderate mean-level change in self-esteem, such as socioeconomic status, social relationships, and life events. However, since in the present meta-analysis these factors cannot be examined (because information on these characteristics is not available for the majority of studies and because measures of these characteristics are often not comparable across studies), here we only briefly summarize the findings (for reviews, see Orth, 2017b; Orth & Robins, in press; Trzesniewski et al., 2013). Research generally suggests that individuals with high socioeconomic status (as indicated by level of education, income, or occupational prestige) have higher self-esteem at each age than individuals with low socioeconomic status (the effect size is of small to medium size), but that the pattern of mean-level change holds across different levels of socioeconomic status (Orth et al., 2012; Orth et al., 2010; Wagner, Gerstorf, et al., 2013; Wagner et al., 2014). Moreover, the evidence suggests that, across the life course, social relationships influence self-esteem (Gruenenfelder-Steiger, Harris, & Fend, 2016; Harris et al., 2017; Orth, 2018; Srivastava & Beer, 2005). For example, using propensity score matching, two longitudinal studies found that romantic relationship transitions such as beginning a relationship influenced people's level of self-esteem (Luciano & Orth, 2017; Wagner, Becker, Lüdtke, & Trautwein, 2015). Finally, longitudinal studies suggest that stressful life events such as becoming unemployed or contracting a chronic disease have the potential to alter people's self-esteem trajectory (Orth & Luciano, 2015; Pettit & Joiner, 2001; Tetzner, Becker, & Baumert, 2016).

Unresolved Issues

Despite the evidence reviewed above, there is a need to better understand the factors that account for individual differences in self-esteem development across the life span. For example, little evidence is available about the moderating effect of cultural context. Theory suggests that cultures may influence the typical level of self-esteem among their members (Heine, Lehman, Markus, & Kitayama, 1999; Markus & Kitayama, 1991). However, according to cross-sectional data, the general pattern of age differences in self-esteem holds across many countries (Bleidorn et al., 2016; Robins et al., 2002). Nevertheless, given the paucity of longitudinal research on self-esteem development outside of Western countries, it is unknown whether individuals from different cultural contexts differ in the normative trajectory of self-esteem over the life span.

Birth cohort is another important factor that might moderate self-esteem development. In other words, do members of different generations show the same life-span trajectory of self-esteem? In fact, it is crucial to test for cohort differences in the life-span trajectory of any construct because if cohort differences are present, then insights from longitudinal studies must be qualified by noting the specific generation for which they are valid. Several studies suggest that members of more recent generations have higher self-esteem and experience steeper increases in self-esteem (Gentile, Twenge, & Campbell, 2010; Twenge & Campbell, 2001; Twenge, Carter, & Campbell, 2017). In fact, in Western countries, sociocultural changes during the past decades may have influenced young people's self-esteem (Gentile et al., 2010). For example, self-esteem has become a more prominent topic in educational contexts and in the media, leading parents, caregivers, and teachers to focus more strongly on the promotion of children's and adolescents' self-esteem. Since these sociocultural changes also may have caused increases in young people's level of narcissism (for the debate, see Trzesniewski & Donnellan, 2010; Twenge & Campbell, 2010), Twenge (2006) has suggested that the cohorts born in the

1970s to 1990s should be called "Generation Me." With regard to self-esteem, however, the findings from several studies conflict with the notion that there have been generational increases (Erol & Orth, 2011; Hamamura & Septarini, 2017; Orth et al., 2015; Trzesniewski & Donnellan, 2010). In particular, evidence from cohort-sequential longitudinal studies suggests that people's average level and slope of the life-span trajectory have not significantly changed over the past generations (Orth et al., 2015; Orth et al., 2012; Orth et al., 2010; but see Twenge et al., 2017).

The present meta-analysis may contribute to knowledge about moderating factors of mean-level change in self-esteem, by testing whether sample characteristics account for between-study variability in effect sizes. In addition to country and birth cohort, we will test for the effects of gender, ethnicity, sample type, and measure of self-esteem.

The Need for the Present Meta-Analysis

Mean-level change in self-esteem has been summarized in a previous meta-analysis (Huang, 2010) that included data from 59 studies with 130 independent samples. Relatively few effect sizes were available for samples above age 22 years. Here, we briefly summarize the findings. There was a small increase in self-esteem in samples younger than 12 years, no significant change in samples aged 12–18, a stronger increase in samples aged 18–30, and no significant change after age 30. Moderator analyses were conducted only for the age group 12–18 years, suggesting that effect sizes were moderated by birth year (with a negative effect on mean-level change) and measure of self-esteem, but not by gender.

However, several methodological and substantive reasons strongly suggest that the present meta-analysis is needed to better understand self-esteem development across the life span. A first important reason is that the effect size measure used by Huang (2010) did not account for the time interval across which the effect was observed. More precisely, Huang

(2010) meta-analyzed *d* values (Cohen, 1988), regardless of whether the amount of change was observed across, e.g., one year, two years, or 10 years. However, when the goal is to estimate the magnitude of change at a given age, meaningful effect size measures must take the length of the time interval into account (i.e., the interval between the two assessments for which the mean change *d* is computed). As an example, if one study yields $d = 0.20$ and another study $d = 0.60$, it is not meaningful to simply average these coefficients if the first effect was observed across a one-year interval and the second across a five-year interval. Thus, instead of using simple coefficients of mean-level change as effect size measure, we suggest that a *change-to-time ratio* should be used. Specifically, in the present meta-analysis we used a coefficient that captures standardized mean-level change per year (denoted as d_{year}). We note that using *d* values and testing whether the time interval moderates the effects is not an appropriate solution to this issue, even if in a given meta-analysis the moderator effect of time interval is nonsignificant. The reason is that the time interval is linked to the amount of change by definition. Thus, we argue that accounting for the time interval must be an integral part of the effect size measure. More detailed information on the effect size measure will be provided in the Method section.

A second reason is that the meta-analysis by Huang (2010) used a restrictive set of inclusion criteria. For example, samples from outside the U.S. were excluded although testing for differences between countries is a key goal in research on self-esteem development. Moreover, samples that were based on participants from a single ethnic group (e.g., African Americans) were excluded; however, including these samples is necessary to gain information about the possible moderating effects of ethnicity. Also, even if the data indicate that ethnicity does not moderate the self-esteem trajectory, including samples from different ethnic groups strongly increases the number of samples and, consequently, the validity and precision of the overall findings.

A third important reason is that since 2007—the most recent publication year of studies included in Huang's (2010) meta-analytic dataset—a large number of new empirical studies on self-esteem development have accumulated that were not available for inclusion in Huang's meta-analysis. For example, as reviewed above, longitudinal research on the life-span trajectory of self-esteem was not available before 2010. Moreover, many longitudinal studies focusing on specific developmental periods such as adolescence, young adulthood, and old age have appeared only after the publication of Huang's meta-analysis. Over the past decade, self-esteem development has become an extensively studied topic in social-personality and developmental psychology, which is reflected by the fact that 70% of the articles considered for inclusion in the present meta-analysis were published in 2008 or later (for more information on the search for studies, see the Method section).

The inclusion of a much larger number of samples (331 in the present research vs. 130 in Huang, 2010) has major implications for the analyses. In the present research, we were able to use much finer-grained age categories than was possible in Huang (2010). In particular, the present meta-analytic dataset allowed us to examine late childhood and adolescence at high temporal resolution (i.e., in 1-year age groups). This is an important advantage because, as reviewed above, inconsistent findings have plagued research on these developmental periods, leading researchers to debate whether self-esteem decreases, remains stable, or even increases in middle childhood and in the transition from childhood to adolescence. Moreover, the present meta-analysis allowed us to examine self-esteem change among older adults with much more power and precision than was possible in the previous meta-analysis. Whereas in Huang (2010), four effect sizes for samples 60 years and older were available, the present dataset included 41 effect sizes for samples over age 60. Thus, the present research enables us to address the question of whether self-esteem remains stable in old age or whether, and to which degree, self-esteem declines in this age period.

Also, the larger number of samples provided much larger statistical power in testing for moderators. Whereas in Huang (2010), moderator tests were possible only for the age group 12–18 years, in the present research we test for moderators across the whole life span. Also, we will test for the moderating effects of sample characteristics that could not be tested by Huang (2010), such as country and ethnicity. The moderator tests are crucial for evaluating whether the meta-analytic findings are robust and can be generalized across differences in gender, ethnicity, country, and birth cohort. Moreover, the larger number of samples significantly reduces the influence of any single study, increasing the validity and robustness of the overall findings. In sum, the larger number of samples in the present meta-analysis allowed us to address important unresolved issues in the field of self-esteem development and to draw a much more precise picture of the life-span trajectory of self-esteem.

Summary of the Goals of the Present Research

The first goal was to comprehensively synthesize the available longitudinal data on mean-level change in self-esteem, to gain precise and robust insights into the normative pattern of the self-esteem trajectory from childhood to old age. At which ages does self-esteem typically increase or decline? More precisely, does self-esteem decline in middle childhood and early adolescence? What is the magnitude of the self-esteem increase during adulthood? Does self-esteem decline in old age and, if so, how strong is the old-age decline? At which point of the life course does self-esteem reach its peak? The second goal was to test for moderators of mean-level change, allowing to evaluate the robustness of the findings. Specifically, we tested for the effects of gender, country, ethnicity, sample type, birth year, and measure of self-esteem.

Method

The present meta-analysis used anonymized data and therefore was exempt from approval by the Ethics Committee of the authors' institution (University of Bern, Faculty of Human Sciences), in accordance with national law.

Selection of Studies

To search for relevant studies we used three strategies. First, English-language journal articles, books, book chapters, and dissertations were searched in the database PsycINFO. We used the following search terms: *self-esteem, self-worth, self-concept, self-liking, self-respect, self-regard, self-acceptance, self-view*, and self-image**. The asterisk (i.e., the truncation symbol) allowed for the inclusion of alternate word endings of the search term (e.g., *self-view** yielded entries containing the term "self-view" but also "self-views"). We restricted the search to empirical-quantitative and longitudinal studies, by using the limitation options "empirical study," "quantitative study," and "longitudinal study" in PsycINFO. This search yielded 1,651 potentially relevant articles, including 125 dissertations. Second, we examined the references cited in four narrative reviews of research on self-esteem development (Orth, 2017b; Orth & Robins, 2014; Robins & Trzesniewski, 2005; Trzesniewski et al., 2013) and cited in three meta-analyses using longitudinal data on self-esteem (Huang, 2010; Sowislo & Orth, 2013; Trzesniewski et al., 2003). This search resulted in 77 additional potentially relevant articles, including 13 dissertations.

To decide on the eligibility of studies, all articles were assessed in full text by the second or third author of this meta-analysis.[1] In addition, a random sample of 60 studies were rated by both the second and third author to obtain estimates of interrater agreement. The interrater agreement on inclusion or exclusion in the meta-analysis was high ($\kappa = .95$) and all diverging assessments were discussed until consensus was reached.

We included dissertations in the meta-analysis because dissertations are a category of the “gray” literature, providing a promising way to examine publication bias (Ferguson & Brannick, 2012; B. D. McLeod & Weisz, 2004). Although dissertations are publicly available and indexed in databases, Ferguson and Brannick (2012) argue that publication bias is less of an issue in dissertations because dissertations are typically submitted to dissertation committees regardless of whether the findings are statistically significant or not. Of course, dissertations are not free from selective reporting, but the argument—which is supported by empirical findings of Ferguson and Brannick (2012) as described below—is that selectivity bias is relatively small in dissertations. Moreover, because dissertations are indexed in databases, meta-analysts can conduct an exhaustive search of this type of study and avoid selection bias in sampling relevant dissertations. In contrast, Ferguson and Brannick (2012) argue that it is almost impossible to obtain a truly random, nonselective sample of unpublished manuscripts by making announcements to electronic mailing lists or writing to researchers working in the field. Consistent with this reasoning, Ferguson and Brannick (2012) reported that unpublished manuscripts typically yield effect sizes that are closer to effect sizes from peer-reviewed journal articles than effect sizes from dissertations. Thus, with regard to the strategy of soliciting unpublished studies, Ferguson and Brannick (2012) suggest that this strategy may even “result in a sample of studies that is *more* rather than *less* biased than the population of published studies” (p. 121, italics in the quote as in the original).

Studies were included in the meta-analysis if the following criteria were fulfilled: (a) self-esteem was assessed using an explicit measure of global self-esteem; (b) the study used a longitudinal study design (i.e., it included two or more assessments of the same sample); (c) the time lag between the first and last assessment was 6 months or longer (note that if a study included more than two assessments, each interval coded was at least 6 months or longer and intervals coded did not overlap); (d) the measure of self-esteem was identical across assessments (i.e., with regard to number of items, item wording, response scale, etc.); (e) the sample included at least 30 participants; (f) the sample was not a clinical sample; (g) the sample as a whole did not undergo a psychological or psychopharmacological intervention (i.e., the sample was not a treatment group of an intervention study; however, we used information from control groups if the control group did not undergo any alternative treatment); and (h) enough information was given to compute effect sizes.

Moreover, studies were included only if (i) the sample was sufficiently homogeneous with regard to age, as operationalized by a cutoff value of *SD* = 5 years for age at Time 1. This inclusion criterion is needed to ensure that the study can provide a valid estimate of age-related change in self-esteem. If age variability in a sample is high, it is unclear whether the observed change in self-esteem can be validly related to the average age in the sample. In particular, if the age distribution of a sample covers developmental periods with distinct patterns of normative change in self-esteem, then the average change in self-esteem could be a misleading estimate of normative change at the average age of the sample. Thus, if all samples included in the meta-analysis had no, or very small, variability of age, all effect sizes could be mapped with high precision on age. On the other hand, it is important to accept some degree of age variability in the samples for several reasons. If we used a strict exclusion criterion for age variability, we would ignore a large number of studies. However, many studies, particularly in adulthood, use samples that are heterogeneous in age. By excluding these studies we would significantly decrease the statistical power of the meta-analysis. Moreover, ignoring a large number of studies could lead to a selective meta-analytic sample because studies with more versus less age variability might also differ in other sample characteristics. Thus, there is a trade-off with regard to being strict versus liberal in including samples with age variability. Given that the literature does not suggest any cutoff value for age variability, we decided to use the value noted above (i.e., *SD* = 5 years). This value was derived rationally from examining the range of standard deviations in samples that seemed acceptable in terms of age variability. When describing the meta-analytic dataset in the Results section, we will examine the distribution of this sample characteristic. We expected that for most samples age variability was much smaller than the cutoff value.

These procedures left 191 articles for analysis, providing effect sizes on 331 independent samples.

Coding of Studies

We coded the following data: year of publication, publication type, sample size, sample type, proportion of female participants, country in which sample was collected, ethnicity, year of Time 1 assessment, measure of self-esteem, mean age of participants at Time 1, standard deviation of age at Time 1, time lag between assessments, and effect size information.

If studies provided information that allowed coding independent subsamples (e.g., female and male participants; U.S. sample versus Chinese sample), we coded subsamples rather than the full sample because this increases the precision of moderator analyses. If year of Time 1 assessment was not reported in the article or in other publications or sources of information on the sample, we estimated it using the following formula: Year of Time 1 assessment = publication year − 3 years (assuming that studies were published on average 3 years after the completion of data collection) − interval between first and last assessment (i.e., duration of data collection). If studies did not report the mean age of participants but valid indicators of age were given, we used this information to estimate age. For example, if a study reported that participants were children in 5th grade, we estimated mean age of participants as 11 years (thus, the general rule was adding the value of 6 to the grade). To take another example, if a study examined a sample of undergraduate students but did not report the mean age, we estimated it to be 20 years (as done by, e.g., Starr & Davila, 2008).

As reported above, studies were excluded if the standard deviation of Time 1 age was larger than 5 years (i.e., if the age variability in the sample was too large). Some studies did not report the standard deviation of age, although all other information needed for including the study was available. We included these studies if other information clearly suggested that the sample was sufficiently homogeneous with regard to age (e.g., if all participants were children in the same grade). Moreover, if a study included a sample with a broad age range (e.g., from young adulthood to old age) and the raw data were available, we coded the study separately for sufficiently homogeneous age groups. In these cases, we used the following procedure to create age groups: We rounded age to full years, started with the youngest age available in the sample, and computed consecutive 5-year age groups across the observed age range of the sample. Corresponding to the inclusion criteria, we computed effect sizes only for those age groups that included at least 30 participants.

For studies that included more than two assessments, we coded all available assessments if the intervals between assessments were 6 months or longer. Later, in the meta-analytic computations, we ensured that each study provided only one effect size estimate per analysis. Thus, when a study provided more than one effect size for a given age period to be meta-analyzed (e.g., age 20 to 25 years), we first averaged effect sizes within studies (by computing the mean) and then conducted the meta-analytic computations. If a study included more than two assessments, but the intervals between assessment were shorter than 6 months, we used those assessments that provided for consecutive (i.e., non-overlapping) intervals that were at least 6 months long. For example, if a study included 5 assessments with 3-month-intervals, we used the first, third, and fifth assessment to compute effect sizes that were based on 6-month-intervals. For studies that provided more than one effect size (e.g., because two different measures of self-esteem were used), we averaged the effect sizes within studies before the meta-analytic computations (by computing the mean) to ensure the statistical independence of effect sizes.

As effect size measure, we used the standardized mean change *d* per year (denoted as d_{year}). We first computed the standardized mean change by subtracting the Time 1 mean in self-esteem from the Time 2 mean in self-esteem and dividing this difference by the Time 1 standard deviation of self-esteem, following the procedures used in the meta-analysis on mean-level change in personality by Roberts et al. (2006; see also Morris & DeShon, 2002). Thus, computing standardized mean changes yielded *d* values (Cohen, 1988), with positive *d* values indicating an increase in self-esteem and negative *d* values indicating a decrease. Next, we set the standardized mean change in relation to the observed time interval, by dividing it by the length of the time lag between Time 1 and Time 2. Thus, the effect size measure used in the present meta-analysis is a change-to-time ratio, with the unit *d* per year. If information on the means and standard deviation of self-esteem was not given in the article, but *d* values of mean-level change were reported, we used these to compute d_{year}.[2]

The articles were coded by the second or third author of this meta-analysis, except for four articles that were coded by the first author. A random sample of 40 studies were coded by both the second and third author to obtain estimates of interrater agreement. The interrater agreement was high ($\kappa \geq .97$ for categorical variables and $r \geq .97$ for continuous variables), except for one variable (see below). All diverging assessments were discussed until consensus was reached. For the categories of sample type, as used in coding the studies, interrater agreement was much lower ($\kappa = .73$). Inspection of the crosstabulation showed that the disagreement resulted from overlap between two categories, specifically "community samples (convenience)" and "community samples (regionally representative);" regionally representative was defined as representative for a region such as a county or city. Given the overlap, we merged these two categories into one category (denoted as community sample), resulting in high agreement for the revised variable of sample type ($\kappa = 1.00$). The revised variable included the following categories: nationally representative, community, and college students.

Meta-Analytic Procedure

The meta-analytic computations were made with R (R Core Team, 2017), using the metafor package (Viechtbauer, 2010). In the effect size analyses, we used random-effects models (for estimating weighted mean effect sizes) and mixed-effects meta-regression models (for testing moderators), following recommendations by Borenstein, Hedges, Higgins, and Rothstein (2009) and Raudenbush (2009). Between-study heterogeneity (i.e., τ^2) was estimated with the DerSimonian–Laird method (DerSimonian & Laird, 1986; Viechtbauer, 2010). Following Borenstein et al. (2009), study weights are given by

$$\omega_i = \left(\frac{1}{v_i + \tau^2}\right),$$

where ω_i is the study weight for study *i*, v_i is the within-study variance for study *i*, and τ^2 is the estimate of between-study heterogeneity. When using standardized mean change as effect size, the within-study variance is given by

$$v_i = \frac{2(1 - r_i)}{n_i} + \frac{{d_i}^2}{2n_i},$$

where d_i is the effect size in study *i*, n_i is the sample size in study *i*, and r_i is the correlation between pre- and post-scores in study *i* (Becker, 1988; Lipsey & Wilson, 2001). Because this correlation is frequently not reported in primary studies, we used an estimate of *r*, as suggested by the methodological literature (Borenstein et al., 2009; Lipsey & Wilson, 2001; Morris & DeShon, 2002). A previous meta-analysis estimated the mean test-retest correlation of self-esteem as .50, based on data from 168 longitudinal studies (Trzesniewski et al., 2003). In the present research, we used this estimate in computing the sampling variance of the effect sizes.

For the effect size analyses, we did not use the mean age at Time 1 as age variable but, instead, the mean age at the center of the time interval on which the effect size was based. For example, if a sample was assessed at age 20 years at Time 1 and age 24 years at Time 2, the age at the center of the interval on which the effect size was based was 22 years. Although the difference between the two age variables (i.e., age at Time 1 and age at the center of the interval) may be irrelevant for short intervals (e.g., one year), the difference is more relevant for long intervals (e.g., 10 years). Because mean-level change in self-esteem might change systematically across long intervals (e.g., the slope might become smaller or larger with age), the most meaningful age value related to the observed effect size is the center of the Time 1–Time 2 interval rather than age at the beginning (Time 1) or end (Time 2) of the interval. Thus, we used the mean age at the center of the time interval for constructing age groups (for further information on age groups, see below) and it was possible that the beginning and the end of an interval fell into separate age groups. We note that it would be ideal if all relevant studies had used relatively short intervals between assessments such as one or two years, which would increase the precision with which effect sizes are mapped on age. However, some studies did use longer intervals and it was important to include them in the meta-analysis to comprehensively summarize all available data. If we excluded these studies, this would decrease the power of the meta-analysis and, moreover, this could lead to a selective meta-analytic dataset because studies with long versus short intervals might also differ in other sample characteristics. Specifically, in 25% of the studies the Time 1–Time 2 interval was longer than 2 years and in 1% longer than 10 years. Thus, a relevant number of studies used medium-sized intervals. However, the number of studies with very long intervals (i.e., longer than 10 years) was very small, suggesting that in this meta-analytic dataset, effect sizes can be mapped with sufficient precision on age.

Results

Description of Studies

The meta-analytic dataset included 331 samples (Table 1 shows basic sample characteristics). Data were drawn from 175 journal articles, 15 dissertations, and 1 book chapter. These 191 articles were published between 1975 and 2016, with the median in 2009. Sample sizes ranged from 32 to 13,401 (M = 498.1, SD = 935.1, Mdn = 236.0). In sum, the samples included 164,868 participants. Eighty-six percent of the samples were community samples, 8% were samples of college students, and 6% were nationally representative. The mean proportion of female participants was 53% (range = 0% to 100%, SD = 34%, Mdn = 52%). Sixty-one percent of the samples were from the United States, 8% from Germany, 4% from Australia, 4% from Canada, 4% from the Netherlands, 4% from Switzerland, 3% from China, 2% from Belgium, 2% from Finland, 2% from the United Kingdom, and the remaining 6% from other countries. Taken together, most samples (96%) were from Western cultural contexts such as the United States, European countries, Australia, and Canada; the remaining samples (4%) were from East and Southeast Asian countries including China, Indonesia, Korea, and Taiwan; no African, South American, or Central American samples were included. With regard to ethnicity, 56% of the samples were predominantly White/European ("predominantly" was defined as 80% and more), 4% predominantly Asian, 2% predominantly Black, 2% predominantly Hispanic/Latin American, 2% predominantly Native American, and 34% were other/mixed. Mean age at Time 1 ranged from 4.3 to 86.6 years (M = 21.9, SD = 16.9; note that some studies included 3 or more waves of data, and that we used the mean age at the center of time intervals for the effect size analyses, so the largest mean age examined was 93.7 years). Year of Time 1 assessment ranged from 1966 to 2011 (M = 1995.2, SD = 9.6). We computed mean year of birth using the variables mean age at Time 1 and year of Time 1 assessment. Mean year of birth ranged from 1899 to 2002 (M = 1973.3, SD = 21.2). To assess self-esteem, 61% of the studies used the Rosenberg Self-Esteem Scale (Rosenberg, 1965), 22% one of the scales by Harter (e.g., Harter, 2012b), 4% one of the scales by Marsh (e.g., Marsh, 1990), and 13% another measure (for overviews of measures of self-esteem, see Blascovich & Tomaka, 1991; Donnellan, Trzesniewski, & Robins, 2015).

As reported in the Method section, one inclusion criterion for studies was that the sample was sufficiently homogeneous with regard to the age of participants (using a cutoff value of SD = 5 years for age at Time 1). This criterion is needed to ensure that effect sizes can be mapped with sufficient precision on age. Across samples, the standard deviation of age was relatively small with a mean of 0.72 years. For 98% of the samples, the standard deviation was not larger than 2 years; 1% of the values were between 2 and 3 years; and 1% were between 3 and 5 years. Supplemental Figure S1 shows a scatterplot of the standard deviation by mean age of the sample. Samples with larger standard deviations (i.e., more than 2 years) mostly included adults. Larger age variability might be more of a concern in child samples than in adult samples, because the precise age might have stronger implications in childhood than in adulthood. In sum, these findings suggest that age variability in the samples was not a concern in this meta-analysis.

As also reported above, some studies provided effect sizes for more than one age. The reason is that some of the longitudinal studies included more than two waves of data, allowing to compute effect sizes for more than one interval. Specifically, the number of intervals ranged from 1 to 10 across studies. Because the goal of the meta-analysis was to comprehensively summarize all available data on mean-level change in self-esteem, it was important not to ignore information that multi-wave studies provided at later waves (i.e., Waves 3 and later). In particular, in studies using samples from middle adulthood and old age, later waves provided valuable information because the number of samples was lower for these developmental periods compared to adolescence and young adulthood. With regard to these multi-wave studies, the following two procedures should be noted. First, as described earlier, if a study included more than two assessments, each interval coded was at least 6 months or longer (as also required for 2-wave studies) and intervals coded from the same study did not overlap. Second, we ensured that all meta-analytic computations were conducted with independent samples (i.e., no participant provided information for more than one effect size included in the same analysis). Therefore, for each of the analyses, we first averaged effect sizes within studies and then conducted the meta-analytic computations. For this reason, we had to use different datasets depending on the specific analysis. For example, in the moderator analyses for age 4–94 years, we used information from all 331 samples that provided effect sizes, by first averaging effect sizes within studies. In contrast, in the effect size analyses, which were conducted within age groups, we averaged effect sizes from multi-wave studies only within the specific age group. The data can be accessed at https://osf.io/zmn5h/?view_only=6eb18c9aebad42ce80a3274ef3d20f51.

Preliminary Analyses

The distribution of effect sizes suggested that there were two effect sizes with large positive values that qualified as potential outliers. When we examined these values formally using the “influence” command of the metafor package (Viechtbauer, 2010), the results suggested that the effect sizes were not influential, as indicated by their DFFITS values (i.e., difference between the predicted average effect for the study with vs. without including it in model fitting; Viechtbauer & Cheung, 2010) and τ^2_{del} values (i.e., decrease in the amount of heterogeneity when study is removed from the meta-analytic dataset; Viechtbauer & Cheung, 2010). The sample sizes of the two studies were small (i.e., 38 and 56, respectively), congruent with the fact that outliers are more likely in small samples, but due to their small size the samples were not influential. The mean age in these samples was 48 and 13 years, respectively, which suggests that the potential outliers were not systematically related to age. We therefore retained the two studies in the meta-analytic dataset, consistent with methodological literature advising against routine deletion of studies with particularly large or small effect sizes (Viechtbauer & Cheung, 2010).

Then, we assessed whether there was evidence of publication bias in the data. We expected that publication bias would not be a problem in this meta-analysis because many studies included did not focus on self-esteem development (i.e., they examined other research questions), but simply reported the relevant statistics (i.e., means and standard deviations of self-esteem) together with statistics on a larger set of variables. We used three methods to test for publication bias. First, we examined the funnel graph, which displays the relation between effect size and the inverse standard error of the effect size (Sutton, 2009). The funnel graph exhibited a symmetrical shape typical of nonbiased meta-analytic datasets (Figure 1). Second, Egger's regression test (Egger, Smith, Schneider, & Minder, 1997) was nonsignificant, $z = 1.42$, $p = .156$, suggesting that the funnel graph did not deviate significantly from a symmetrical shape (the test was conducted with metafor; Viechtbauer, 2010). Third, we compared effect sizes from dissertations (as a category of gray literature) with effect sizes from peer-reviewed journal articles. If dissertations yield effect sizes that differ significantly from journal articles, this is evidence of publication bias. All effect sizes from dissertations ($k = 28$) were based on samples from childhood to beginning adulthood; specifically, the mean age of the samples ranged from 7 to 19 years. We therefore restricted the test to this age range ($k = 193$ for journal articles). The results of a mixed-effects meta-regression model indicated that effect sizes from dissertations and journal articles did not differ significantly, $z = 1.40$, $p = .161$. Thus, the three methods converged in suggesting that there was no evidence of publication bias.

Effect Size Analyses

To gain an overview of the effect size data, we first inspected a scatterplot of the relation between effect size and age (Figure 2). The figure shows that the variability of effect sizes was much larger in childhood and adolescence than in adulthood.[3] Moreover, the scatterplot suggested that the majority of effect sizes were positive during childhood and adolescence (indicating an increase in self-esteem during these developmental periods) and that effect sizes gradually declined across adulthood and tended to be below zero in old age.

As the goal of this meta-analysis was to map mean-level change in self-esteem on age, effect size analyses were conducted within age groups. For these analyses, we constructed multiple age groups across the observed age range (see Table 2). For the age range from 10 to 20 years, the meta-analytic dataset included a large number of samples. Given the substantive importance of this developmental period, we constructed 1-year age groups. Because the numbers of studies were lower for other developmental stages, we constructed 2-year age groups from 4 to 10 years, 5-year age groups from 20 to 30 years, and 10-year age groups from 30 to 90 years. The oldest age group included effect sizes ranging from age 90 to 94 years.

Although the power of significance tests of mean-level change would be greater if we constructed broader age groups (which then would include a larger number of samples), it is important to emphasize that null-hypothesis significance testing of mean-level change was not a central goal in this meta-analysis (cf. Cumming, 2014; Fraley & Marks, 2007; Greenwald, 1975). As reviewed in the Introduction, prior research has provided sufficient evidence suggesting that self-esteem does change significantly across the life course. In the present research, the goal was rather to obtain estimates of age-dependent mean-level change and, thus, narrower age groups provide more precision with regard to age. We used the weighted mean effect size (i.e., the point estimate) as best estimate of mean-level change in the age group, regardless of whether the estimate differed significantly from zero or not. In Table 2, we report the null-hypothesis significance tests of mean effect sizes for reasons of completeness.

Table 2 reports the meta-analytic findings for all age groups from 4 to 94 years. Most of the weighted mean effect sizes had a positive sign, except for two years in early adolescence and for the age groups of 70 years and older. The largest yearly mean-level change in self-esteem emerged for early childhood (age 4–6), age 10, and for the years from age 15 to 29 years (all with a positive sign), as well as for age 90 to 94 years (with a negative sign).[4]

Figure 3 summarizes the findings by aggregating the point estimates of mean-level change across the observed age range from 4 to 94 years. The vertical axis shows cumulative *d* values (cumulative *d*s are relative to age 4; for a similar way to illustrate meta-analytic findings on mean-level change, see Roberts et al., 2006). For age groups that covered more than one year (e.g., age 4–6; age 20–25), the estimate of yearly change (i.e., d_{year}) was used for each year included in the age group. The figure shows that, on average, self-esteem increased during early and middle childhood (i.e., until age 11; cumulative $d = 0.34$). Between age 11 and 15 (cumulative $d = 0.34$), self-esteem remained constant. Afterwards, self-esteem increased strongly until age 30 (cumulative $d = 1.05$; thus, self-esteem increased by $d = 0.71$ from age 15 to 30). Self-esteem increased slightly from age 30 to 60 (cumulative $d = 1.30$; the difference between age 30 and 60 corresponded to a small effect size of $d = 0.25$), peaking at age 60 years and remaining constant until age 70. Then, self-esteem declined slightly until age 90 (cumulative $d = 1.15$), but more strongly from age 90 to 94 (cumulative $d = 0.76$).

The findings on heterogeneity reported in Table 2 correspond well to the inconsistencies in findings reported in the literature on self-esteem development in middle childhood and adolescence. For these developmental periods, the I^2 values (i.e., the ratio of total heterogeneity by total variability) were large. Nevertheless, at the same time, the total sample sizes were also large and the confidence intervals of the meta-analytic effect sizes were relatively small (at least from age 10), suggesting that the weighted mean effect sizes provide reliable estimates of mean-level change in middle childhood and adolescence.

Moderator Analyses

As noted above, the analyses on mean effect sizes showed that there was significant variation of effect sizes. We therefore tested whether sample characteristics moderated the effect sizes. The variables mean year of birth and proportion of female participants were continuous and were included as such in the moderator variables. For the categorical variables, we focused on specific contrasts due to low numbers of samples in some of the categories. For country, we contrasted samples from the United States (61%) with samples from other countries (39%). For ethnicity, we contrasted samples that were White/European (56%) with other samples (44%). For measure of self-esteem, we contrasted samples that were assessed with the Rosenberg scale (61%) with samples that were assessed with other measures (39%). Finally, for sample type, we focused on the contrast between nationally representative (6%) and other samples (94%), because representative samples provide more valid results compared with nonrepresentative samples.

We first examined the full sample of studies covering the observed age range from 4 to 94 years. Table 3 shows the intercorrelations among the effect size and moderators (see the values below the diagonal). Given that many of the intercorrelations were substantial, it is possible that the relations between the moderators and effect sizes are confounded if the moderator effects are not mutually controlled for each other. Consequently, it was important to use multiple regression analysis to gain information about the independent effects of moderators on effect sizes. Table 4 shows the results of the mixed-effects meta-regression models. In the moderator analyses, we also controlled for mean age of the sample. In these analyses, it was sufficient to control for the linear effect of age on effect size. Given that the effect size (i.e., yearly change in self-esteem) represents the slope of the trajectory, a linear age effect (e.g., a linear age-dependent decline of the slope) captures a quadratic trajectory, corresponding roughly to the shape of mean-level change shown in Figure 3.[5]

Table 4 shows that, when all moderators were mutually controlled for in the regression analysis, linear age had the expected negative effect on yearly change in self-esteem. Moreover, the measure of self-esteem (i.e., Rosenberg scale vs. other measures) was positively related to effect size. However, no other variable had a significant effect on the effect sizes. Thus, the findings do not suggest that sample characteristics such as mean year of birth, gender, country, and ethnicity influence the average shape of the life-span trajectory of self-esteem. Moreover, effect sizes did not differ significantly between nationally representative and other samples, which strengthens the generalizability of the findings (of note, in the multiple regression analysis, the regression coefficient of sample type was close to zero). It should also be noted that the regression coefficient of year of birth was negative (albeit nonsignificant). Thus, if anything, more recent generations showed less positive slopes of self-esteem change compared to older generations, which is opposite to the predictions from the Generation Me hypothesis (Twenge, 2006).

Given that the majority of samples (i.e., 220 of 331 samples) had a mean age ranging from 10 to 20 years, we repeated the moderator analyses for this subsample of studies. Because these samples were relatively homogeneous with regard to age, these analyses may yield additional insights into the effects of moderators. In particular, holding age (relatively) constant may provide a more valid test of cohort effects on mean-level change (Baltes et al., 1979). Table 3 (values above the diagonal) shows the intercorrelations among the effect size and moderators and Table 4 (values in the right half) shows the multiple regression coefficients of the moderators. Except for the age effect, which was nonsignificant for this set of studies, the findings were essentially the same as for the full set of studies. Measure of self-esteem was positively related to effect size. All other moderators did not show significant effects. Again, year of birth had a negative (albeit nonsignificant) effect on mean-level change in self-esteem, although the Generation Me hypothesis would predict a positive effect.

Two of the predictors—more precisely, mean age and mean year of birth—showed a very large correlation in the full meta-analytic sample (−.92), which raises concerns related to multicollinearity. On the one hand, two variables that are so strongly correlated should typically not simultaneously be included in a multiple regression model. On the other hand, in this particular case, it is essential to mutually control the variables and to test their unique effects for substantive reasons. As reviewed in the Introduction, an important goal of research on self-esteem development is to disentangle true developmental effects (i.e., as indicated by age) and cohort effects (i.e., as indicated by birth year). Thus, if only one of the variables is included in the model, then it is possible that its observed effect is significantly confounded for the very reason that age and birth year are substantially correlated. We therefore included both variables in the meta-regression models. However, given concerns about multicollinearity, we replicated the models by omitting one of the variables, specifically mean year of birth (Supplemental Table S1). The results showed that the pattern of findings was virtually unaltered, supporting the conclusions from the models including both variables. Moreover, we note that in the subset of samples with mean age ranging from 10 to 20 years, the correlation between age and birth year was much lower (−.34) than in the full meta-analytic dataset. The fact that the pattern of findings was similar in this subset of samples further reduces concerns about multicollinearity and strengthens confidence in the conclusions from the moderator analyses.

The central conclusion from the moderator analyses is that the findings on mean-level change in self-esteem do not significantly differ for samples from different birth cohorts, samples from different countries, samples with different compositions in terms of gender and ethnicity, and for different sample types.

Discussion

In this meta-analysis, we synthesized the available longitudinal data on mean-level change in self-esteem. The analyses were based on 331 independent samples, including data from 164,868 participants. The mean age associated with the effect sizes ranged from 4 to 94 years. The results show that people's self-esteem changes systematically across the life span. Average levels of self-esteem increased from age 4 to 11, remained stable from age 11 to 15, then increased strongly until age 30 and more slowly until age 60, peaked between age 60 and 70 years, and declined after age 70. The cumulative mean-level change from early childhood to the peak at the end of middle adulthood corresponded to a very large effect size. Moderator analyses were conducted for the full set of samples and for the subset of samples between ages 10 and 20 years. Although the measure of self-esteem accounted for differences in effect sizes, the moderator analyses suggested that the pattern of mean-level change held in samples from different birth cohorts and different countries, samples with different compositions in terms of gender and ethnicity, and for different sample types.

Implications of the Findings

Childhood. As noted in the Introduction, narrative reviews of the literature had suggested that average levels of self-esteem decrease in the transition from early to middle childhood (i.e., from about age 4 to 8 years; Harter, 2006c, 2012a), due to the effects of social-cognitive advances, such as the ability to distinguish between the actual and ideal self and the ability to use social comparison information (Harter, 2006c; Ruble et al., 1980; Stipek & Tannatt, 1984). These effects have been described as ironic because although children make progress in their social-cognitive development in this developmental period, these very advances were assumed to have negative side effects on children's self-esteem. It should be noted, however, that in this age group most of the relevant studies did not use measures of global self-esteem but assessed self-evaluations in specific domains such as academic abilities and peer relationships (J. Eccles et al., 1993; Marsh, 1989; Marsh et al., 1984; Stipek & Tannatt, 1984). Thus, little evidence was available with regard to mean-level change in global self-esteem.

The meta-analytic findings on the self-esteem trajectory in childhood are in strong contrast to the depictions included in narrative reviews, but suggested that there is a positive trajectory in children's global self-esteem already in the transition from early to middle childhood (the cumulative *d* was 0.21 for the period from age 4 to 8 years). Which mechanisms could account for the increase in global self-esteem during childhood? One possible process might be gains in personal autonomy and in the general sense of mastery, a mechanism that has been discussed also with regard to increases in self-esteem in middle and late adolescence (Harter, 2006c). For example, in a longitudinal study of the self-esteem trajectory from age 14 to 30 years, the normative increase in sense of mastery explained a large portion of the normative increase in participants' self-esteem (Erol & Orth, 2011). Nevertheless, we note that future research on the self-esteem trajectory in childhood is needed, given that for age 4–8 years the present meta-analysis included data from only seven samples (fortunately, however, for age 8–10 the number of samples was larger, with 13 samples).

Adolescence. The meta-analytic findings on adolescence also deviate, at least partially, from the typical description in the literature. According to Harter (2006c), self-esteem declines at the transition from childhood to adolescence (at about 11–13 years) and begins to recover in middle adolescence (at about age 15 years). Thus, early adolescents might go through a low point in self-esteem (Harter, 2006c; Robins et al., 2002). In contrast, the present research indicates that self-esteem is, on average, constant from age 11 to 15 years. However, consistent with Harter's (2006c) description, self-esteem began to increase strongly at age 15 years. Thus, although the upward trend in self-esteem, as seen in childhood, is interrupted during early and middle adolescence, the present findings suggest that this developmental period is not a time of "storm and stress" in adolescents' self-concept (Arnett, 1999; Hollenstein & Lougheed, 2013; Molloy, Ram, & Gest, 2011). Nevertheless, it is important to emphasize that the present research describes the normative trajectory and that there are, of course, individual differences in the trajectory. Thus, some adolescents may experience declines in their self-esteem due to pubertal changes, conflicts with parents, and mood disruptions in this developmental period.

Adulthood. The shape of the self-esteem trajectory across adulthood identified in this meta-analysis corresponds well to the general shape of the trajectory found in previous studies, which suggested that self-esteem follows an inverted U-shape over the life course (note that these studies began only in adolescence or young adulthood; Orth et al., 2015; Orth et al., 2012; Orth et al., 2010). Clearly, however, effect sizes differed across prior studies. For example, in Orth et al. (2012), the increase from age 16 (i.e., the youngest age assessed in that study) to the peak at age 51 was relatively small with $d = 0.29$, whereas in other studies as well as in this meta-analysis the effect size was much larger (in this meta-analysis, the cumulative effect for the same age interval was $d = 0.86$). We therefore believe that this meta-analysis significantly advances the field, because it provides a much more precise picture of the life-span trajectory compared to any single primary study. The reason is that the meta-analytic method provides for much larger statistical power (by aggregating across a large number of samples) and also for greater robustness of the findings, because the peculiarities of primary studies typically cancel each other out.

Old age. As noted in the Introduction, previous research had yielded inconsistent findings about the magnitude of the normative decline in self-esteem in old age and, moreover, about the age at which the decline begins (e.g., Orth et al., 2010; von Soest et al., 2017; Wagner, Gerstorf, et al., 2013). The meta-analytic findings suggest self-esteem does decline in old age. However, the rate of change was small until age 90 (the cumulative mean-level change from age 70 to 90 corresponded to $d = -0.15$). Moreover, it should be noted that the decline started at a very high level of self-esteem. In fact, individuals reached their highest level of self-esteem at age 60–70 years. Thus, the present research suggests that many people are able to maintain a relatively high level of self-esteem even during old age.

After age 90 years, however, self-esteem declined more strongly (from age 90 to 94, cumulative $d = -0.39$). Thus, this meta-analysis supports the notion that there is a terminal decline of self-esteem in very old age, with a much larger rate of decline (for terminal decline in life satisfaction, see Gerstorf, Ram, Estabrook, et al., 2008). This issue is important because if very old adults experience significant loss in self-esteem, these changes may impair their level of well-being and contribute to the emergence of depressive symptoms and disorders (Orth, Robins, Trzesniewski, Maes, & Schmitt, 2009; Sowislo & Orth, 2013). However, note that the estimate for the age group 90–94 years was based on only two studies, suggesting that more research is needed to gain more precise insights into the self-esteem trajectory in very old age.

Birth cohort. The moderator analyses indicated that birth cohort did not explain variability in the effect sizes. Given that mean year of birth ranged from 1899 to 2002 across the samples included in this meta-analysis, the nonsignificant cohort effect indicates that the shape of the life-span trajectory has not changed over the generations born during the 20th century. Of course, samples that differed in their mean year of birth typically also differed in their mean age, providing effect sizes for different parts of the life span (although some of the samples from different generations provided information about the same age, because year of data collection also varied across samples). However, the moderator analyses yield valid insights into cohort effects for two reasons. First, the effect of birth year was controlled for the effect of age in the moderator analyses. Thus, the effect of birth year captured the unique cohort effect while holding age constant (and also while controlling for other variables such as ethnicity, which otherwise could have confounded the effect). Second, we repeated the moderator analyses in the subset of samples aged 10 to 20 years. Thus, in this analysis, age varied relatively little across samples, while birth year varied more strongly (for this subset of samples, birth year ranged from 1950 to 2000, with M = 1982.1 and SD = 10.5 years). The fact that the second moderator analysis replicated the findings based on the full set of samples (yielding a nonsignificant cohort effect) corroborates the conclusion that there were no systematic cohort effects on mean-level change in self-esteem.

Consequently, the meta-analytic findings do not support Twenge's (2006) claim that the birth cohorts called Generation Me (i.e., cohorts born in the 1970s to 1990s) experience more positive, steeper increases in self-esteem. Moreover, we note that the sign of the cohort effect was negative in both moderator analyses (albeit nonsignificant), pointing in the direction of less positive mean-level change in samples from more recent generations. Thus, if anything, the pattern of findings was opposite to predictions about Generation Me. Obviously, the present findings do not speak to the debate about generational changes in other personality characteristics such as narcissism, which are not considered here (for the debate see, e.g., Trzesniewski, Donnellan, & Robins, 2008; Twenge, Konrath, Foster, Campbell, & Bushman, 2008).

The nonsignificant cohort effect is relevant also for methodological reasons, because the modeling of a coherent life-span trajectory based on estimates across different birth cohorts (as done in this meta-analysis; see Figure 3) is valid only if the cohorts included do not differ in significant ways in age-related mean-level change (i.e., the estimates for specific ages must not differ across cohorts). For example, if more recent generations experienced steeper increases in self-esteem during adolescence and young adulthood compared to older generations, these cohort differences would confound the overarching life-span trajectory, leading to wrong conclusions about the patterns of self-esteem development. Thus, the fact that in this meta-analysis cohort effects were nonsignificant supports the validity of conclusions about the shape of the life-span trajectory.

It is important to note that this meta-analysis provides information only about cohort effects on mean-level change or, in other words, about cohort effects on the slope but not the overall level of the trajectory. Thus, other evidence is needed to evaluate the hypothesis that more recent generations show higher levels of self-esteem. As reviewed in the Introduction, cohort-sequential longitudinal studies are particularly suited to assessment of cohort differences in the level and slope of developmental trajectories (Baltes et al., 1979; Duncan et al., 2006). The evidence from several cohort-sequential studies with large samples overall suggests that there are no systematic cohort differences in average levels of self-esteem (Erol & Orth, 2011; Orth et al., 2015; Orth et al., 2012; Orth et al., 2010), even if for one of these datasets cohort effects have been reported (Twenge et al., 2017). Moreover, studies with other research designs have also questioned the claim that more recent generations experience higher levels of self-esteem compared to older generations (Hamamura & Septarini, 2017; Trzesniewski & Donnellan, 2010).

Specificity of the life-span trajectory of self-esteem. Given that self-esteem is substantially correlated with neuroticism, with cross-sectional correlations at about −.60 (Judge, Erez, Bono, & Thoresen, 2002; Robins, Tracy, Trzesniewski, Potter, & Gosling, 2001; Watson et al., 2002), it may be informative to compare the life-span trajectory of self-esteem with neuroticism (or, put differently, emotional stability, i.e., the positive pole of the neuroticism dimension). Many studies indicate that emotional stability increases across large parts of adulthood (Milojev & Sibley, 2017; Roberts et al., 2006; Terracciano et al., 2005; but see Allemand, Zimprich, & Hendriks, 2008; Lucas & Donnellan, 2011). Thus, this pattern corresponds to the uptrend in self-esteem during young and middle adulthood. However, the evidence suggests that emotional stability and self-esteem follow different trajectories in childhood and, possibly, old age. Whereas emotional stability tends to decline in the first decade of life (Soto, 2016; Wängqvist, Lamb, Frisén, & Hwang, 2015), the present research suggests that self-esteem increases from age 4 to 10 years. Also, the available evidence suggests that emotional stability further increases, or at least is relatively stable, during old age (Lucas & Donnellan, 2011; Milojev & Sibley, 2017; Roberts et al., 2006; Specht et al., 2011; Terracciano et al., 2005), whereas this meta-analysis suggests that self-esteem declines in old age, with a sharper drop in very old age. The divergent developmental patterns for self-esteem and neuroticism support that it is essential to distinguish between the constructs and to understand the specific trajectory of self-esteem across the life span.

Limitations and Strengths

As is true for each meta-analysis, it is important to evaluate whether there is evidence of publication bias in the data (Sutton, 2009). If publication bias were present, estimates of mean-level change could be distorted, e.g., towards more positive effect sizes. However, we argue that the a priori likelihood of publication bias in the present data is low, given that the raw statistics that we used for computing the effect sizes were often not a central focus of the primary studies (i.e., they were included in tables with descriptive statistics on the measures used). Moreover, as reported in the Results section, the funnel graph of effect sizes exhibited an inconspicuous pattern. Also, Egger's test of funnel graph asymmetry was nonsignificant (Egger et al., 1997). Finally, effect sizes did not differ significantly between articles in peer-reviewed journals and dissertations. Dissertations are a category of the gray literature recommended for estimating effect sizes in unpublished studies because they are indexed in databases and can be searched systematically (Ferguson & Brannick, 2012; B. D. McLeod & Weisz, 2004). Moreover, dissertations provide for a stricter test of publication bias compared to unpublished manuscripts, because dissertations typically yield effect sizes that differ more strongly from peer-reviewed journal articles than do effect sizes from unpublished manuscripts (Ferguson & Brannick, 2012). Thus, the empirical analyses strongly suggest that the present findings are not biased by selective publishing.

Another potential threat to the validity of the present findings is selectivity in the samples used in primary studies. For example, if most primary studies used samples of participants with particularly high (or low) levels of well-being or particularly favorable (or adverse) environmental conditions, estimates of mean-level change in self-esteem could be biased. However, the present findings speak against this possibility, because effect sizes did not differ between nationally representative and other types of samples (our dataset included 21 nationally representative samples with a total of more than 25,000 individuals). Given that sample type was not a significant moderator, this suggests that the full meta-analytic sample of studies allows for valid insights into the normative pattern of mean-level change in self-esteem across the life span.

With regard to ethnicity, we had to focus on the contrast between samples that were White/European versus other samples, due to the very low proportion of samples that consisted predominantly of Black, Asian, or Hispanic individuals. Thus, although the present meta-analysis suggests that individuals who are European or of European descent show a trajectory that corresponds to the trajectory identified for the full set of samples, conclusions about the trajectory among other ethnic groups are based on an omnibus test and should be considered tentative. Given this limitation, researchers interested in the self-esteem trajectory among specific ethnic groups should consult relevant primary studies (Gray-Little & Hafdahl, 2000; Orth et al., 2010; Robins et al., 2002; Shaw et al., 2010).

Few samples were from outside of Western cultural contexts (only 4% of the samples were from Asian countries and no samples were available for African, South American, or Central American countries). Although it would have been desirable to test whether the life-span trajectory of self-esteem holds outside of Western countries, this was not possible in this meta-analysis. In future longitudinal studies, it would be highly interesting to test whether the trajectory identified in the present research replicates in non-Western samples (for cross-sectional studies, see Bleidorn et al., 2016; Robins et al., 2002). Nevertheless, the moderator analyses support the conclusion that the trajectory holds across Western countries.

The moderator analyses indicated that the measure of self-esteem explained variability in the effect sizes. Specifically, samples that were assessed with the Rosenberg Self-Esteem Scale (Rosenberg, 1965) showed greater effect sizes than samples that were assessed with other measures. This effect held for both the full set of samples and for the reduced set of samples aged 10 to 20 years. The reason for this effect is unclear. One possibility is that the Rosenberg scale shows greater reliability than the other measures of self-esteem and, if so, then the larger amount of measurement error in other measures could dampen estimates of mean change more strongly. Thus, future research could test whether differences in the reliability of self-esteem measures contribute to differences in estimates of mean-level change. Nevertheless, research suggests that not only the Rosenberg scale but also the Harter scales and Marsh scales—together, these scales were used in 87% of the samples included in the present meta-analysis—generally have good reliability (for a review, see Donnellan et al., 2015). Moreover, it may be useful to note a large body of research supports the validity of the Rosenberg, Harter, and Marsh scales (Donnellan et al., 2015). Also, these measures show other favorable psychometric properties such as measurement invariance across time (Marsh, Scalas, & Nagengast, 2010; Orth, Erol, Ledermann, & Grob, 2018; Orth, Robins, Widaman, & Conger, 2014), age (Orth et al., 2015; Whiteside-Mansell & Corwyn, 2003), gender (Byrne, 1988), and ethnicity and cultural contexts (Alessandri, Vecchione, Eisenberg, & Łaguna, 2015; Bodkin-Andrews, Ha, Craven, & Yeung, 2010; Leung, Marsh, Craven, & Abduljabbar, 2016). For these reasons, we believe that the moderator effect of measure of self-esteem does not significantly threaten the validity of the overall findings of the meta-analysis.

A limitation is that we did not test whether the weighted mean effect sizes differed significantly between age groups. The reason is that the age groups were not fully independent from each other, because some multi-wave studies (i.e., studies with three waves and more) provided effect sizes for more than one age group. However, it is important to emphasize that in the analyses within age groups, as well as in the moderator analyses, independence of samples was ensured.

As reported in the Method section, when studies provided more than one effect size because two different measures of self-esteem were used or because for a given age group more than one time interval was available, we ensured that each sample provided only one effect size per analysis by averaging effect sizes within samples before the meta-analytic computations. It should be noted, however, that it would be statistically superior to use a multilevel meta-analytic approach to account for dependent effect size estimates. Also, we note that it is possible that the computation of the sampling variance of standardized mean change would need an adjustment when using d_{year} instead of d. Given that the meta-analytic literature seems to be mute on this issue, we used the equation for the sampling variance as described in Becker (1988) and Lipsey and Wilson (2001). A finding, for which the explanation is unclear, is that for age groups in middle adulthood and old age, the Q heterogeneity statistics were small when compared to their corresponding degrees of freedom (under homogeneity, the expected value for Q is $df = k – 1$; Borenstein et al., 2009). Nevertheless, it should be noted that prior research on self-esteem development in adulthood, at least until about age 70, reported relatively consistent findings, corresponding to the low degree of heterogeneity.

The use of age groups allowed to estimate age-dependent change in self-esteem in a relatively precise way (i.e., we distinguished between 22 age groups). Although it is often preferable to examine age as a continuous variable, in the present context this would have required to test specific models of change across the life span (e.g., to test whether the trajectory can be captured by a linear, quadratic, or cubic model). In fact, the meta-analytic findings, as shown in Figure 3, correspond roughly—but not fully—to the quadratic life-span trajectory suggested by several longitudinal studies (Orth et al., 2015; Orth et al., 2012; Orth et al., 2010). Importantly, however, imposing a change model on the present meta-analytic data would not have allowed to estimate change in a fine-grained way in developmental periods such as early adolescence and old age. For example, it is likely that the discontinuous shape of the trajectory between age 11 and 15 years would not have been captured when modeling a coherent trajectory across the life span. Yet, as discussed above, the precise estimates for age groups in childhood, adolescence, and old age allowed for important conclusions with regard to previously unresolved issues in the field of self-esteem development.

An important strength of this meta-analysis is the comprehensive coverage of the available data across the life span (i.e., from as young as age 4 years to 94 years). Also, we believe that the effect size measure used in the present research (i.e., the standardized mean change *d* per year, d_{year}, which is a change-to-time ratio) significantly contributes to the validity of the findings, compared to the simpler effect size measures (such as the simple *d*) used in prior meta-analyses of mean-level change in personality constructs. The reason is that the length of the time interval across which change is observed should be an integral part of the effect size measure. Based on age-specific estimates of d_{year}, the meta-analysis allowed drawing a precise picture of the normative trajectory from early childhood to old age, by showing cumulative *d* values across the life span.

Conclusions

Based on longitudinal data from 331 independent samples with more than 160,000 individuals, this meta-analysis shows that people’s self-esteem changes in a systematic, normative way over the life span. The findings suggest that, on average, self-esteem increases in early and middle childhood, remains constant in adolescence, increases strongly in young adulthood, continues to increase in middle adulthood, peaks between age 60 and 70 years, and declines in old age and more strongly in very old age. The pattern of results did not differ significantly for samples from different birth cohorts and different countries, samples with different compositions in terms of gender and ethnicity, and for different sample types, suggesting that the findings are robust and generalizable within Western cultural contexts. With regard to childhood and early adolescence, the meta-analytic findings deviate from prior depictions found in the literature, which had been based on narrative reviews. Thus, the notion that self-esteem declines in middle childhood and reaches a low point in early adolescence should be revised. Moreover, although there is a normative decline in self-esteem after age 70, the meta-analysis indicates that the magnitude of decline is relatively benign until age 90, suggesting that many older adults are able to maintain a relatively high level of self-esteem. In very old age, however, self-esteem declines more strongly, suggesting that in this developmental stage individuals may be vulnerable due to declining self-esteem and that interventions aimed at improving self-esteem may be needed.

References

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

Ackerman, R. A., Witt, E. A., Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., & Kashy, D. A. (2011). What does the Narcissistic Personality Inventory really measure? *Assessment, 18*, 67-87. http://dx.doi.org/10.1177/1073191110382845

*Adachi, P. J. C., & Willoughby, T. (2014). It's not how much you play, but how much you enjoy the game: The longitudinal associations between adolescents' self-esteem and the frequency versus enjoyment of involvement in sports. *Journal of Youth and Adolescence, 43*, 137-145. http://dx.doi.org/10.1007/s10964-013-9988-3

Alessandri, G., Vecchione, M., Eisenberg, N., & Łaguna, M. (2015). On the factor structure of the Rosenberg (1965) General Self-Esteem Scale. *Psychological Assessment, 27*, 621-635. http://dx.doi.org/10.1037/pas0000073

Allemand, M., Zimprich, D., & Hendriks, A. A. J. (2008). Age differences in five personality domains across the life span. *Developmental Psychology, 44*, 758-770. http://dx.doi.org/10.1037/0012-1649.44.3.758

*Antonishak, J. M. (2005). Predictors and sequelae of instability in adolescent peer groups. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 66*(3-B), 1752.

Anusic, I., & Schimmack, U. (2016). Stability and change of personality traits, self-esteem, and well-being: Introducing the meta-analytic stability and change model of retest correlations. *Journal of Personality and Social Psychology, 110*, 766-781. http://dx.doi.org/10.1037/pspp0000066

Arnett, J. J. (1999). Adolescent storm and stress, reconsidered. *American Psychologist, 54*, 317-326. http://dx.doi.org/10.1037/0003-066X.54.5.317

*Arunkumar, R., Midgley, C., & Urdan, T. (1999). Perceiving high or low home–school dissonance: Longitudinal effects on adolescent emotional and academic well-being. *Journal of Research on Adolescence, 9*, 441-466. http://dx.doi.org/10.1207/s15327795jra0904_4

*Au, R. C. P., Watkins, D. A., & Hattie, J. A. C. (2010). Academic risk factors and deficits of learned hopelessness: A longitudinal study of Hong Kong secondary school students. *Educational Psychology, 30*, 125-138. http://dx.doi.org/10.1080/01443410903476400

*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. http://dx.doi.org/10.1037/0022-3514.35.6.365

Baird, B. M., Lucas, R. E., & Donnellan, M. B. (2010). Life satisfaction across the lifespan: Findings from two nationally representative panel studies. *Social Indicators Research, 99*, 183-203. http://dx.doi.org/10.1007/s11205-010-9584-9

*Baldwin, S. A., & Hoffmann, J. P. (2002). The dynamics of self-esteem: A growth-curve analysis. *Journal of Youth and Adolescence, 31*, 101-113. http://dx.doi.org/10.1023/A:1014065825598

Baltes, P. B., Cornelius, S. W., & Nesselroade, J. R. (1979). Cohort effects in developmental psychology. In J. R. Nesselroade & P. B. Baltes (Eds.), *Longitudinal research in the study of behavior and development* (pp. 61-87). New York, NY: Academic Press.

Baltes, P. B., & Mayer, K. U. (1999). *The Berlin Aging Study: Aging from 70 to 100*. Cambridge, England: Cambridge University Press.

*Bao, X., & Jin, K. (2015). The beneficial effect of Tai Chi on self-concept in adolescents. *International Journal of Psychology, 50*, 101-105. http://dx.doi.org/10.1002/ijop.12066

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest, 4,* 1-44. http://dx.doi.org/10.1111/1529-1006.01431

Becker, B. J. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology, 41*, 257-278. http://dx.doi.org/10.1111/j.2044-8317.1988.tb00901.x

*Becker, M., Cortina, K. S., Tsai, Y. M., & Eccles, J. S. (2014). Sexual orientation, psychological well-being, and mental health: A longitudinal analysis from adolescence to young adulthood. *Psychology of Sexual Orientation and Gender Diversity, 1*, 132-145. http://dx.doi.org/10.1037/sgd0000038

*Birkeland, M. S., Melkevik, O., Holsen, I., & Wold, B. (2012). Trajectories of global self-esteem development during adolescence. *Journal of Adolescence, 35*, 43-54. http://dx.doi.org/10.1016/j.adolescence.2011.06.006

Blascovich, J., & Tomaka, J. (1991). Measures of self-esteem. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 115-160). San Diego, CA: Academic Press.

*Blatny, M., Millova, K., Jelinek, M., & Osecka, T. (2015). Personality predictors of successful development: Toddler temperament and adolescent personality traits predict well-being and career stability in middle adulthood. *Plos One, 10*, e0126032. http://dx.doi.org/10.1371/journal.pone.0126032

Bleidorn, W., Arslan, R. C., Denissen, J. J. A., Rentfrow, P. J., Gebauer, J. E., Potter, J., & Gosling, S. D. (2016). Age and gender differences in self-esteem: A cross-cultural window. *Journal of Personality and Social Psychology, 111*, 396-410. http://dx.doi.org/10.1037/pspp0000078

Bleidorn, W., Hufer, A., Kandler, C., Hopwood, C. J., & Riemann, R. (2018). A nuclear twin family study of self-esteem. *European Journal of Personality*, Advance online publication. http://dx.doi.org/10.1002/per.2136

Blyth, D. A., Simmons, R. G., & Carlton-Ford, S. (1983). The adjustment of early adolescents to school transitions. *Journal of Early Adolescence, 3*, 105-120. http://dx.doi.org/10.1177/027243168331008

Bodkin-Andrews, G. H., Ha, M. T., Craven, R. G., & Yeung, A. S. (2010). Factorial invariance testing and latent mean differences for the Self-Description Questionnaire II (Short Version) with indigenous and non-indigenous Australian secondary school students. *International Journal of Testing, 9*, 47-79. http://dx.doi.org/10.1080/15305050903352065

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.

*Bosacki, S. L. (2015). Children's theory of mind, self-perceptions, and peer relations: A longitudinal study. *Infant and Child Development*, *24*, 175-188. http://dx.doi.org/10.1002/icd.1878

Bosson, J. K., Lakey, C. E., Campbell, W. K., Zeigler-Hill, V., Jordan, C. H., & Kernis, M. H. (2008). Untangling the links between narcissism and self-esteem: A theoretical and empirical review. *Social and Personality Psychology Compass, 2*, 1415-1439. http://dx.doi.org/10.1111/j.1751-9004.2008.00089.x

*Brenning, K., Soenens, B., Van Petegem, S., & Vansteenkiste, M. (2015). Perceived maternal autonomy support and early adolescent emotion regulation: A longitudinal study. *Social Development, 24,* 561-578. http://dx.doi.org/10.1111/sode.12107

*Brown, K. M., McMahon, R. P., Biro, F. M., Crawford, P., Schreiber, G. B., Similo, S. L., . . . Striegel-Moore, R. (1998). Changes in self-esteem in Black and White girls between the ages of 9 and 14 years. *Journal of Adolescent Health, 23*, 7-19. http://dx.doi.org/10.1016/S1054-139X(97)00238-3

*Brummelman, E., Thomaes, S., Nelemans, S. A., Orobio de Castro, B., Overbeek, G., & Bushman, B. J. (2015). Origins of narcissism in children. *Proceedings of the National Academy of Sciences, 112*, 3659-3662. http://dx.doi.org/10.1073/pnas.1420870112

Brummelman, E., Thomaes, S., & Sedikides, C. (2016). Separating narcissism from self-esteem. *Current Directions in Psychological Science, 25*, 8-13. http://dx.doi.org/10.1177/0963721415619737

*Burke, P. J., & Harrod, M. M. (2005). Too much of a good thing? *Social Psychology Quarterly, 68,* 359–374. http://dx.doi.org/10.1177/019027250506800404

*Burwell, R. A., & Shirk, S. R. (2006). Self processes in adolescent depression: The role of self-worth contingencies. *Journal of Research on Adolescence, 16,* 479-490. http://dx.doi.org/10.1111/j.1532-7795.2006.00503.x

Byrne, B. M. (1988). Measuring adolescent self-concept: Factorial validity and equivalency of the SDQ III across gender. *Multivariate Behavioral Research, 23*, 361-375. http://dx.doi.org/10.1207/s15327906mbr2303_5

*Cairns, E., McWhirter, L., Duffy, U., & Barry, R. (1990). The stability of self-concept in late adolescence: Gender and situational effects. *Personality and Individual Differences, 11*, 937-944. http://dx.doi.org/10.1016/0191-8869(90)90275-V

*Cantin, S., & Boivin, M. (2004). Change and stability in children's social network and self-perceptions during transition from elementary to junior high school. *International Journal of Behavioral Development, 28*, 561-570. http://dx.doi.org/10.1080/01650250444000289

*Caprara, G. V., Alessandri, G., Barbaranelli, C., & Vecchione, M. (2013). The longitudinal relations between self-esteem and affective self-regulatory efficacy. *Journal of Research in Personality, 47*, 859-870. http://dx.doi.org/10.1016/j.jrp.2013.08.011

Carstensen, L. L., Pasupathi, M., Mayr, U., & Nesselroade, J. R. (2000). Emotional experience in everyday life across the adult life span. *Journal of Personality and Social Psychology, 79*, 644-655. http://dx.doi.org/10.1037/0022-3514.79.4.644

Charles, S. T., Reynolds, C. A., & Gatz, M. (2001). Age-related differences and change in positive and negative affect over 23 years. *Journal of Personality and Social Psychology, 80*, 136-151. http://dx.doi.org/10.1037/0022-3514.80.1.136

*Chen, X., He, Y., & Li, D. (2004). Self-perceptions of social competence and self-worth in Chinese children: Relations with social and school performance. *Social Development, 13*, 570-589. http://dx.doi.org/10.1111/j.1467-9507.2004.00284.x

Chung, J. M., Hutteman, R., van Aken, M. A. G., & Denissen, J. J. A. (2017). High, low, and in between: Self-esteem development from middle childhood to young adulthood. *Journal of Research in Personality, 70*, 122-133. http://dx.doi.org/10.1016/j.jrp.2017.07.001

*Chung, J. M., Robins, R. W., Trzesniewski, K. H., Noftle, E. E., Roberts, B. W., & Widaman, K. F. (2014). Continuity and change in self-esteem during emerging adulthood. *Journal of Personality and Social Psychology, 106*, 469-483. http://dx.doi.org/10.1037/a0035135

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159. http://dx.doi.org/10.1037/0033-2909.112.1.155

*Colarossi, L. G., & Eccles, J. S. (2003). Differential effects of support providers on adolescents' mental health. *Social Work Research, 27,* 19-30. http://dx.doi.org/10.1093/swr/27.1.19

Cole, D. A., Maxwell, S. E., Martin, J. M., Peeke, L. G., Seroczynski, A. D., Tram, J. M., . . . Maschman, T. (2001). The development of multiple domains of child and adolescent self-concept: A cohort sequential longitudinal design. *Child Development, 72*, 1723-1746. http://dx.doi.org/10.1111/1467-8624.00375

Coleman, P. G., Ivani-Chalian, C., & Robinson, M. (1993). Self-esteem and its sources: Stability and change in later life. *Ageing and Society, 13*, 171-192. http://dx.doi.org/10.1017/S0144686x00000842

Crocker, J., & Luhtanen, R. K. (2003). Level of self-esteem and contingencies of self-worth: Unique effects on academic, social, and financial problems in college students. *Personality and Social Psychology Bulletin, 29*, 701-712. http://dx.doi.org/10.1177/0146167203029006003

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review, 108*, 593-623. http://dx.doi.org/10.1037/0033-295X.108.3.593

*Crocker, P. R. E., Sabiston, C. M., Kowalski, K. C., McDonough, M. H., & Kowalski, N. (2006). Longitudinal assessment of the relationship between physical self-concept and health-related behavior and emotion in adolescent girls. *Journal of Applied Sport Psychology, 18*, 185-200. http://dx.doi.org/10.1080/10413200600830257

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*, 7-29. http://dx.doi.org/10.1177/0956797613504966

*Davison, K. K., Schmalz, D. L., Young, L. M., & Birch, L. L. (2008). Overweight girls who internalize fat stereotypes report low psychosocial well-being. *Obesity, 16*, S30-S38. http://dx.doi.org/10.1038/oby.2008.451

*Davison, K. K., Werder, J. L., Trost, S. G., Baker, B. L., & Birch, L. L. (2007). Why are early maturing girls less active? Links between pubertal development, psychological well-being, and physical activity among girls at ages 11 and 13. *Social Science and Medicine, 64*, 2391-2404. http://dx.doi.org/10.1016/j.socscimed.2007.02.033

*de Araujo, P., & Lagos, S. (2013). Self-esteem, education, and wages revisited. *Journal of Economic Psychology, 34*, 120-132. http://dx.doi.org/10.1016/j.joep.2012.12.001

*DeRosier, M. E. (2004). Building relationships and combating bullying: Effectiveness of a school-based social skills group intervention. *Journal of Clinical Child and Adolescent Psychology, 33*, 196-201. http://dx.doi.org/10.1207/S15374424JCCP3301_18

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*, 177-188. http://dx.doi.org/10.1016/0197-2456(86)90046-2

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin, 125*, 276-302. http://dx.doi.org/10.1037/0033-2909.125.2.276

*Dohnt, H., & Tiggemann, M. (2006). The contribution of peer and media influences to the development of body satisfaction and self-esteem in young girls: A prospective study. *Developmental Psychology, 42,* 929-936. http://dx.doi.org/10.1037/0012-1649.42.5.929

Donnellan, M. B., Kenny, D. A., Trzesniewski, K. H., Lucas, R. E., & Conger, R. D. (2012). Using trait-state models to evaluate the longitudinal consistency of global self-esteem from adolescence to adulthood. *Journal of Research in Personality, 46*, 634-645. http://dx.doi.org/10.1016/j.jrp.2012.07.005

*Donnellan, M. B., Trzesniewski, K. H., Conger, K. J., & Conger, R. D. (2007). A three-wave longitudinal study of self-evaluations during young adulthood. *Journal of Research in Personality*, *41*, 453-472. http://dx.doi.org/10.1016/j.jrp.2006.06.004

Donnellan, M. B., Trzesniewski, K. H., & Robins, R. W. (2011). Self-esteem: Enduring issues and controversies. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences* (pp. 718-746). Chichester, UK: Wiley-Blackwell. http://dx.doi.org/10.1002/9781444343120.ch28

Donnellan, M. B., Trzesniewski, K. H., & Robins, R. W. (2015). Measures of self-esteem. In G. J. Boyle, D. H. Saklofske, & G. Matthews (Eds.), *Measures of personality and social psychological constructs* (pp. 131-157). London, UK: Elsevier. http://dx.doi.org/10.1016/B978-0-12-386915-9.00006-1

Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., Moffitt, T. E., & Caspi, A. (2005). Low self-esteem is related to aggression, antisocial behavior, and delinquency. *Psychological Science, 16,* 328-335. http://dx.doi.org/10.1111/j.0956-7976.2005.01535.x

*Dorn, L. D., Susman, E. J., & Ponirakis, A. (2003). Pubertal timing and adolescent adjustment and behavior: Conclusions vary by rater. *Journal of Youth and Adolescence, 32*, 157-167. http://dx.doi.org/10.1023/A:1022590818839

*Duckworth, A. L., Tsukayama, E., & May, H. (2010). Establishing causality using longitudinal hierarchical linear modeling: An illustration predicting achievement from self-control. *Social Psychological and Personality Science, 1*, 311-317. http://dx.doi.org/10.1177/1948550609359707

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2006). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. Mahwah, NJ: Erlbaum.

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. http://dx.doi.org/10.1111/j.1467-8624.1993.tb02946.x

Eccles, J. S., Wigfield, A., Flanagan, C. A., Miller, C., Reuman, D. A., & Yee, D. (1989). Self-concepts, domain values, and self-esteem: Relations and changes at early adolescence. *Journal of Personality, 57*, 283-310. http://dx.doi.org/10.1111/j.1467-6494.1989.tb00484.x

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629-634. http://dx.doi.org/10.1136/bmj.315.7109.629

*Eisenberg, M. E., Neumark-Sztainer, D., Haines, J., & Wall, M. (2006). Weight-teasing and emotional well-being in adolescents: Longitudinal findings from Project EAT. *Journal of Adolescent Health, 38*, 675-683. http://dx.doi.org/10.1016/j.jadohealth.2005.07.002

Erol, R. Y., & Orth, U. (2011). Self-esteem development from age 14 to 30 years: A longitudinal study. *Journal of Personality and Social Psychology, 101*, 607-619. http://dx.doi.org/10.1037/a0024299

*Erol, R. Y., & Orth, U. (2014). Development of self-esteem and relationship satisfaction in couples: Two longitudinal studies. *Developmental Psychology, 50*, 2291-2303. http://dx.doi.org/10.1037/a0037370

*Falci, C. D. (2006). Gender and race trajectories of adolescent depressed mood: The dynamic role of social support, the self concept and stress exposure. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 67*(7-A), 2757.

*Fay, K. E. (2011). Weighing in on the issue: A longitudinal analysis of the individual and contextual factors that influence the developmental trajectories of eating disorders among adolescents. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 72*(9-B), 5599.

*Feldman, G. S. (1999). Self-efficacy and outcome expectancies in successful student development: A multi-ethnic study of the college transition. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 60*(6-B), 2939.

*Fenzel, L. M. (2000). Prospective study of changes in global self-worth and strain during the transition to middle school. *Journal of Early Adolescence, 20*, 93-116. http://dx.doi.org/10.1177/0272431600020001005

Ferguson, C. J., & Brannick, M. T. (2012). Publication bias in psychological science: Prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychological Methods, 17*, 120-128. http://dx.doi.org/10.1037/a0024445

*Flynn, M., Cicchetti, D., & Rogosch, F. (2014). The prospective contribution of childhood maltreatment to low self-worth, low relationship quality, and symptomatology across adolescence: A developmental-organizational perspective. *Developmental Psychology, 50,* 2165-2175. http://dx.doi.org/10.1037/a0037162

*Foynes, M. M., Smith, B. N., & Shipherd, J. C. (2015). Associations between race-based and sex-based discrimination, health, and functioning: A longitudinal study of Marines. *Medical Care, 53*, S128-S135. http://dx.doi.org/10.1097/MLR.0000000000000300

Fraley, R. C., & Marks, M. J. (2007). The null hypothesis significance-testing debate and its implications for personality research. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 149-169). New York, NY: Guilford.

*Fredricks, J. A., & Eccles, J. S. (2008). Participation in extracurricular activities in the middle school years: Are there developmental benefits for African American and European American youth? *Journal of Youth and Adolescence, 37*, 1029-1043. http://dx.doi.org/10.1007/s10964-008-9309-4

*Frijns, T., & Finkenauer, C. (2009). Longitudinal associations between keeping a secret and psychosocial adjustment in adolescence. *International Journal of Behavioral Development, 33*, 145-154. http://dx.doi.org/10.1177/0165025408098020

*Galambos, N. L., Barker, E. T., & Krahn, H. J. (2006). Depression, self-esteem, and anger in emerging adulthood: Seven-year trajectories. *Developmental Psychology, 42*, 350-365. http://dx.doi.org/10.1037/0012-1649.42.2.350

*Galliher, R. V., Jones, M. D., & Dahl, A. (2011). Concurrent and longitudinal effects of ethnic identity and experiences of discrimination on psychosocial adjustment of Navajo adolescents. *Developmental Psychology, 47*, 509-526. http://dx.doi.org/10.1037/a0021061

*Gayman, M. D., Lloyd, D. A., & Ueno, K. (2011). The history and timing of depression onset as predictors of young-adult self-esteem. *Journal of Research on Adolescence*, *21*, 691-702. http://dx.doi.org/10.1111/j.1532-7795.2010.00702.x

Gentile, B., Twenge, J. M., & Campbell, W. K. (2010). Birth cohort differences in self-esteem, 1988-2008: A cross-temporal meta-analysis. *Review of General Psychology, 14*, 261-268. http://dx.doi.org/10.1037/a0019919

Gerstorf, D., Ram, N., Estabrook, R., Schupp, J., Wagner, G. G., & Lindenberger, U. (2008). Life satisfaction shows terminal decline in old age: Longitudinal evidence from the German Socio-Economic Panel Study (SOEP). *Developmental Psychology, 44*, 1148-1159. http://dx.doi.org/10.1037/0012-1649.44.4.1148

Gerstorf, D., Ram, N., Röcke, C., Lindenberger, U., & Smith, J. (2008). Decline in life satisfaction in old age: Longitudinal evidence for links to distance-to-death. *Psychology and Aging, 23*, 154-168. http://dx.doi.org/10.1037/0882-7974.23.1.154

*Gest, S. D., Domitrovich, C. E., & Welsh, J. A. (2005). Peer academic reputation in elementary school: Associations with changes in self-concept and academic skills. *Journal of Educational Psychology, 97*, 337-346. http://dx.doi.org/10.1037/0022-0663.97.3.337

*Graham, D. J., Bauer, K. W., Friend, S., Barr-Anderson, D. J., & Neumark-Sztainer, D. (2014). Personal, behavioral, and socioenvironmental correlates of physical activity among adolescent girls: Cross-sectional and longitudinal associations. *Journal of Physical Activity and Health, 11*, 51-61. http://dx.doi.org/10.1123/jpah.2011-0239

Gray-Little, B., & Hafdahl, A. R. (2000). Factors influencing racial comparisons of self-esteem: A quantitative review. *Psychological Bulletin, 126*, 26-54. http://dx.doi.org/10.1037/0033-2909.126.1.26

*Green, J., Liem, G. A. D., Martin, A. J., Colmar, S., Marsh, H. W., & McInerney, D. (2012). Academic motivation, self-concept, engagement, and performance in high school: Key processes from a longitudinal perspective. *Journal of Adolescence, 35*, 1111-1122. http://dx.doi.org/10.1016/j.adolescence.2012.02.016

*Greene, M. L., & Way, N. (2005). Self-esteem trajectories among ethnic minority adolescents: A growth curve analysis of the patterns and predictors of change. *Journal of Research on Adolescence, 15*, 151-178. http://dx.doi.org/10.1111/j.1532-7795.2005.00090.x

Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin, 82*, 1-20. http://dx.doi.org/10.1037/h0076157

Gruenenfelder-Steiger, A. E., Harris, M. A., & Fend, H. A. (2016). Subjective and objective peer approval evaluations and self-esteem development: A test of reciprocal, prospective, and long-term effects. *Developmental Psychology, 52*, 1563-1577. http://dx.doi.org/10.1037/dev0000147

*Gupta, T., Way, N., McGill, R. K., Hughes, D., Santos, C., Jia, Y., . . . Deng, H. (2013). Gender-typed behaviors in friendships and well-being: A cross-cultural study of Chinese and American boys. *Journal of Research on Adolescence, 23*, 57-68. http://dx.doi.org/10.1111/j.1532-7795.2012.00824.x

Hamamura, T., & Septarini, B. G. (2017). Culture and self-esteem over time: A cross-temporal meta-analysis among Australians, 1978-2014. *Social Psychological and Personality Science, 8*, 904-909. http://dx.doi.org/10.1177/1948550617698205

Harris, M. A., Donnellan, M. B., Guo, J., McAdams, D. P., Garnier-Villarreal, M., & Trzesniewski, K. H. (2017). Parental co-construction of 5- to 13-year-olds' global self-esteem through reminiscing about past events. *Child Development, 88*, 1810-1822. http://dx.doi.org/10.1111/cdev.12944

*Harris, M. A., Gruenenfelder-Steiger, A. E., Ferrer, E., Donnellan, M. B., Allemand, M., Fend, H., . . . Trzesniewski, K. H. (2015). Do parents foster self-esteem? Testing the prospective impact of parent closeness on adolescent self-esteem. *Child Development, 86*, 995-1013. http://dx.doi.org/10.1111/cdev.12356

Harris, M. A., Wetzel, E., Robins, R. W., Donnellan, M. B., & Trzesniewski, K. H. (2018). The development of global and domain self-esteem from ages 10 to 16 for Mexican-origin youth. *International Journal of Behavioral Development, 42*, 4-16. http://dx.doi.org/10.1177/0165025416679744

Harter, S. (2006a). The development of self-esteem. In M. H. Kernis (Ed.), *Self-esteem issues and answers: A sourcebook of current perspectives* (pp. 144-150). New York, NY: Psychology Press.

Harter, S. (2006b). Developmental and individual difference perspectives on self-esteem. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 311-334). Mahwah, NJ: Erlbaum.

Harter, S. (2006c). The self. In N. Eisenberg (Ed.), *Handbook of child psychology (6th ed.). Volume 3: Social, emotional, and personality development* (pp. 505-570). Hoboken, NJ: Wiley.

Harter, S. (2012a). *The construction of the self: Developmental and sociocultural foundations*. New York, NY: Guilford.

Harter, S. (2012b). *Self-Perception Profile for Children: Manual and questionnaires (Grades 3-8)*. Denver, CO: University of Denver.

*Heaven, P., & Ciarrochi, J. (2008). Parental styles, gender and the development of hope and self-esteem. *European Journal of Personality, 22*, 707-724. http://dx.doi.org/10.1002/per.699

Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review, 106*, 766-794. http://dx.doi.org/10.1037/0033-295X.106.4.766

Helson, R., & Soto, C. J. (2005). Up and down in middle age: Monotonic and nonmonotonic changes in roles, status, and personality. *Journal of Personality and Social Psychology, 89,* 194-204. http://dx.doi.org/10.1037/0022-3514.89.2.194

*Hirsch, B. J., & Rapkin, B. D. (1987). The transition to junior high school: A longitudinal study of self-esteem, psychological symptomatology, school life, and social support. *Child Development, 58,* 1235-1243. http://dx.doi.org/10.2307/1130617

*Hoge, D. R., Smit, E. K., & Hanson, S. L. (1990). School experiences predicting changes in self-esteem of sixth- and seventh-grade students. *Journal of Educational Psychology, 82*, 117-127. http://dx.doi.org/10.1037/0022-0663.82.1.117

*Hoglund, C. L. (1995). Longitudinal study of gender differences in global and domain-specific self-esteem in children. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 56*(5-B), 2898.

Hollenstein, T., & Lougheed, J. P. (2013). Beyond storm and stress: Typicality, transactions, timing, and temperament to account for adolescent change. *American Psychologist, 68,* 444-454. http://dx.doi.org/10.1037/a0033586

*Hope, E. C., Chavous, T. M., Jagers, R. J., & Sellers, R. M. (2013). Connecting self-esteem and achievement: Diversity in academic identification and disidentification patterns among Black college students. *American Educational Research Journal, 50*, 1122-1151. http://dx.doi.org/10.3102/0002831213500333

Huang, C. (2010). Mean-level change in self-esteem from childhood through adulthood: Meta-analysis of longitudinal studies. *Review of General Psychology, 14*, 251-260. http://dx.doi.org/10.1037/a0020543

*Hubbs-Tait, L., Osofsky, J. D., Hann, D. M., & Culp, A. M. (1994). Predicting behavior problems and social competence in children of adolescent mothers. *Family Relations, 43*, 439-446. http://dx.doi.org/10.2307/585376

Hutteman, R., Hennecke, M., Orth, U., Reitz, A. K., & Specht, J. (2014). Developmental tasks as a framework to study personality development in adulthood and old age. *European Journal of Personality, 28*, 267-278. http://dx.doi.org/10.1002/per.1959

*Hutteman, R., Nestler, S., Wagner, J., Egloff, B., & Back, M. D. (2015). Wherever I may roam: Processes of self-esteem development from adolescence to emerging adulthood in the context of international student exchange. *Journal of Personality and Social Psychology, 108*, 767-783. http://dx.doi.org/10.1037/pspp0000015

*Impett, E. A., Henson, J. M., Breines, J. G., Schooler, D., & Tolman, D. L. (2011). Embodiment feels better: Girls' body objectification and well-being across adolescence. *Psychology of Women Quarterly, 35*, 46-58. http://dx.doi.org/10.1177/0361684310391641

*Ireson, J., & Hallam, S. (2009). Academic self-concepts in adolescence: Relations with achievement and ability grouping in schools. *Learning and Instruction, 19,* 201-213. http://dx.doi.org/10.1016/j.learninstruc.2008.04.001

*Jackson, L. M., Pratt, M. W., Hunsberger, B., & Pancer, S. M. (2005). Optimism as a mediator of the relation between perceived parental authoritativeness and adjustment among adolescents: Finding the sunny side of the street. *Social Development, 14*, 273-304 http://dx.doi.org/10.1111/j.1467-9507.2005.00302.x

Johnson, M. D., & Galambos, N. L. (2014). Paths to intimate relationship quality from parent-adolescent relations and mental health. *Journal of Marriage and Family, 76*(145-160). http://dx.doi.org/10.1111/jomf.12074

*Johnson, M. K. (2013). Parental financial assistance and young adults' relationships with parents and well-being. *Journal of Marriage and Family, 75*, 713-733. http://dx.doi.org/10.1111/jomf.12029

*Jou, Y. H. (2013). Family relationship quality and its psychological effects among Taiwanese adolescents. In C. C. Yi (Ed.), *The psychological well-being of East Asian youth* (pp. 69-88). Dordrecht, The Netherlands: Springer.

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology, 83*, 693-710. http://dx.doi.org/10.1037/0022-3514.83.3.693

*Kahle, L. R., Kulka, R. A., & Klingel, D. M. (1980). Low adolescent self-esteem leads to multiple interpersonal problems: A test of social-adaptation theory. *Journal of*

Personality and Social Psychology, 39, 496-502. http://dx.doi.org/10.1037/0022-3514.39.3.496

*Kakihara, F., Tilton-Weaver, L., Kerr, M., & Stattin, H. (2010). The relationship of parental control to youth adjustment: Do youths' feelings about their parents play a role? *Journal of Youth and Adolescence, 39*, 1442-1456. http://dx.doi.org/10.1007/s10964-009-9479-8

*Kaplan, H. B. (1975). The self-esteem motive and change in self-attitudes. *Journal of Nervous and Mental Disease, 161*, 265-275. http://dx.doi.org/10.1097/00005053-197510000-00006

Kasen, S., Cohen, P., Chen, H., & Castille, D. (2003). Depression in adult women: Age changes and cohort effects. *American Journal of Public Health, 93*, 2061-2066. http://dx.doi.org/10.2105/AJPH.93.12.2061

*Keel, P. K., Fulkerson, J. A., & Leon, G. R. (1997). Disordered eating precursors in pre- and early adolescent girls and boys. *Journal of Youth and Adolescence, 26*, 203-216. http://dx.doi.org/10.1023/A:1024504615742

Kessler, R. C., Foster, C., Webster, P. S., & House, J. S. (1992). The relationship between age and depressive symptoms in two national surveys. *Psychology and Aging, 7*, 119-126. http://dx.doi.org/10.1037/0882-7974.7.1.119

*Kim, J., & Cicchetti, D. (2009). Mean-level change and intraindividual variability in self-esteem and depression among high-risk children. *International Journal of Behavioral Development, 33*, 202-214. http://dx.doi.org/10.1177/0165025408098021

*Kipp, L. E., & Weiss, M. R. (2015). Social predictors of psychological need satisfaction and well-being among female adolescent gymnasts: A longitudinal analysis. *Sport, Exercise, and Performance Psychology, 4*, 153-169. http://dx.doi.org/10.1037/spy0000033

*Kistner, J., David, C., & Rapper, K. (2007). Self-enhancement of peer acceptance: Implications for children's self-worth and interpersonal functioning. *Social Development, 16*, 24-44. http://dx.doi.org/10.1111/j.1467-9507.2007.00370.x

*Kiviruusu, O., Huurre, T., Aro, H., Marttunen, M., & Haukkala, A. (2015). Self-esteem growth trajectory from adolescence to mid-adulthood and its predictors in adolescence. *Advances in Life Course Research, 23,* 29-43. http://dx.doi.org/10.1016/j.alcr.2014.12.003

*Klima, T., & Repetti, R. L. (2008). Children's peer relations and their psychological adjustment differences between close friendships and the larger peer group. *Merrill-Palmer Quarterly, 54*, 151-178. http://dx.doi.org/10.1353/mpq.2008.0016

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. http://dx.doi.org/10.1037/0033-2909.125.4.470

Krueger, J. I., Vohs, K. D., & Baumeister, R. F. (2008). Is the allure of self-esteem a mirage after all? *American Psychologist, 63,* 64-65. http://dx.doi.org/10.1037/0003-066X.63.1.64

*Kuperminc, G. P., Blatt, S. J., Shahar, G., Henrich, C., & Leadbeater, B. J. (2004). Cultural equivalence and cultural variance in longitudinal associations of young adolescent self-definition and interpersonal relatedness to psychological and school adjustment. *Journal of Youth and Adolescence, 33*, 13-30. http://dx.doi.org/10.1023/A:1027378129042

Kuster, F., & Orth, U. (2013). The long-term stability of self-esteem: Its time-dependent decay and nonzero asymptote. *Personality and Social Psychology Bulletin, 39*, 677-690. http://dx.doi.org/10.1177/0146167213480189

*Kuster, F., Orth, U., & Meier, L. L. (2012). Rumination mediates the prospective effect of low self-esteem on depression: A five-wave longitudinal study. *Personality and Social Psychology Bulletin, 38*, 747-759. http://dx.doi.org/10.1177/0146167212437250

*Kuster, F., Orth, U., & Meier, L. L. (2013). High self-esteem prospectively predicts better work conditions and outcomes. *Social Psychological and Personality Science, 4*, 668-675. http://dx.doi.org/10.1177/1948550613479806

*Kuzucu, Y., Bontempo, D. E., Hofer, S. M., Stallings, M. C., & Piccinin, A. M. (2014). Developmental change and time-specific variation in global and specific aspects of self-concept in adolescence and association with depressive symptoms. *Journal of Early Adolescence, 34*, 638-666. http://dx.doi.org/10.1177/0272431613507498

*LaGrange, B., Cole, D. A., Dallaire, D. H., Ciesla, J. A., Pineda, A. Q., Truss, A. E., & Folmer, A. (2008). Developmental changes in depressive cognitions: A longitudinal evaluation of the Cognitive Triad Inventory for Children. *Psychological Assessment, 20*, 217-226. http://dx.doi.org/10.1037/1040-3590.20.3.217

*Larsen, H., Overbeek, G., Vermulst, A. A., Granic, I., & Engels, R. C. M. E. (2010). Initiation and continuation of best friends and adolescents' alcohol consumption: Do self-esteem and self-control function as moderators? *International Journal of Behavioral Development, 34*, 406-416. http://dx.doi.org/10.1177/0165025409350363

*Le Bars, H., Gernigon, C., & Ninot, G. (2009). Personal and contextual determinants of elite young athletes' persistence or dropping out over time. *Scandinavian Journal of Medicine and Science in Sports, 19*, 274-285. http://dx.doi.org/10.1111/j.1600-0838.2008.00786.x

*Lee, J. (2012). Dynamic growth models of academic achievement and self-esteem in South Korean adolescents. *Social Behavior and Personality, 40*, 1555-1568. http://dx.doi.org/10.2224/sbp.2012.40.9.1555

*Leeuwis, F. H., Koot, H. M., Creemers, D. H. M., & van Lier, P. A. C. (2015). Implicit and explicit self-esteem discrepancies, victimization and the development of late childhood internalizing problems. *Journal of Abnormal Child Psychology, 43*, 909-919. http://dx.doi.org/10.1007/s10802-014-9959-5

*Lemmens, J. S., Valkenburg, P. M., & Peter, J. (2011). Psychosocial causes and consequences of pathological gaming. *Computers in Human Behavior, 27*, 144-152. http://dx.doi.org/10.1016/j.chb.2010.07.015

Leung, K. C., Marsh, H. W., Craven, R. G., & Abduljabbar, A. S. (2016). Measurement invariance of the Self-Description Questionnaire II in a Chinese sample. *European Journal of Psychological Assessment, 32*, 128-139. http://dx.doi.org/10.1027/1015-5759/a000242

*Lin, K. C., Twisk, J. W. R., & Rong, J. R. (2011). Longitudinal interrelationships between frequent geographic relocation and personality development: Results from the Amsterdam Growth and Health Longitudinal Study. *American Journal of Orthopsychiatry, 81*, 285-292. http://dx.doi.org/10.1111/j.1939-0025.2011.01097.x

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lucas, R. E., & Donnellan, M. B. (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology, 101*, 847-861. http://dx.doi.org/10.1037/a0024298

Luciano, E. C., & Orth, U. (2017). Transitions in romantic relationships and development of self-esteem. *Journal of Personality and Social Psychology, 112*, 307-328. http://dx.doi.org/10.1037/pspp0000109

Luhmann, M. (2017). The development of subjective well-being. In J. Specht (Ed.), *Personality development across the lifespan* (pp. 197-218). London, UK: Elsevier. http://dx.doi.org/10.1016/B978-0-12-804674-6.00013-2

*Luyckx, K., Klimstra, T. A., Duriez, B., Van Petegem, S., Beyers, W., Teppers, E., & Gossens, L. (2013). Personal identity processes and self-esteem: Temporal sequences in high school and college students. *Journal of Research in Personality, 47*, 159-170. http://dx.doi.org/10.1016/j.jrp.2012.10.005

MacDonald, G., & Leary, M. R. (2012). Individual differences in self-esteem. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 354-377). New York, NY: Guilford.

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

*Maldonado, L., Huang, Y., Chen, R., Kasen, S., Cohen, P., & Chen, H. (2013). Impact of early adolescent anxiety disorders on self-esteem development from adolescence to young adulthood. *Journal of Adolescent Health, 53*, 287-292. http://dx.doi.org/10.1016/j.jadohealth.2013.02.025

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224-253. http://dx.doi.org/10.1037/0033-295X.98.2.224

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology, 81*, 417-430. http://dx.doi.org/10.1037/0022-0663.81.3.417

Marsh, H. W. (1990). *Self Description Questionnaire (SDQ) II: Manual*. Campbelltown, New South Wales, Australia: University of Western Sydney, SELF Research Centre.

Marsh, H. W., Barnes, J., Cairns, L., & Tidman, M. (1984). Self-Description Questionnaire: Age and sex effects in the structure and level of self-concept for preadolescent children. *Journal of Educational Psychology, 76*, 940-956. http://dx.doi.org/10.1037/0022-0663.76.5.940

*Marsh, H. W., Craven, R., & Debus, R. (1998). Structure, stability, and development of young children's self-concepts: A multicohort-multioccasion study. *Child Development, 69,* 1030-1053. http://dx.doi.org/10.2307/1132361

*Marsh, H. W., Ellis, L. A., Parada, R. H., Richards, G., & Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analysis. *Psychological Assessment, 17*, 81-102. http://dx.doi.org/10.1037/1040-3590.17.1.181

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (2013). Measurement invariance of Big-Five factors over the life span: ESEM tests of gender, age, plasticity, maturity, and la dolce vita effects. *Developmental Psychology, 49,* 1194-1218. http://dx.doi.org/10.1037/a0026913

Marsh, H. W., Scalas, L. F., & Nagengast, B. (2010). Longitudinal tests of competing factor structures for the Rosenberg Self-Esteem Scale: Traits, ephemeral artifacts, and stable response styles. *Psychological Assessment, 22*, 366-381. http://dx.doi.org/10.1037/a0019225

*Marshall, S. L., Parker, P. D., Ciarrochi, J., & Heaven, P. C. L. (2014). Is self-esteem a cause or consequence of social support? A 4-year longitudinal study. *Child Development*, *85*, 1275-1291. http://dx.doi.org/10.1111/cdev.12176

*Martin, A. J., Mansour, M., Anderson, M., Gibson, R., Liem, G. A. D., & Sudmalis, D. (2013). The role of arts participation in students' academic and nonacademic outcomes: A longitudinal study of school, home, and community factors. *Journal of Educational Psychology*, *105*, 709-727. http://dx.doi.org/10.1037/a0032795

*Martins, N., & Harrison, K. (2012). Racial and gender differences in the relationship between children's television use and self-esteem: A longitudinal panel study. *Communication Research, 39*, 338-357. http://dx.doi.org/10.1177/0093650211401376

*Mayzer, R. (2004). First alcohol use and the development of antisocial behavior problems from preschool through early adolescence. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 65*(9-B), 4873.

*McCabe, M. P., Ricciardelli, L. A., & Holt, K. (2005). A longitudinal study to explain strategies to change weight and muscles among normal weight and overweight children. *Appetite, 45*, 225-234. http://dx.doi.org/10.1016/j.appet.2005.07.009

*McCarthy, J. D., & Hoge, D. R. (1982). Analysis of age effects in longitudinal studies of adolescent self-esteem. *Developmental Psychology, 18*, 372-379. http://dx.doi.org/10.1037/0012-1649.18.3.372

*McCarty, C. A., Stoep, A. V., & McCauley, E. (2007). Cognitive features associated with depressive symptoms in adolescence: Directionality and specificity. *Journal of Clinical Child and Adolescent Psychology, 36*, 147-158. http://dx.doi.org/10.1080/15374410701274926

*McGrath, E. P., & Repetti, R. L. (2002). A longitudinal study of children's depressive symptoms, self-perceptions, and cognitive distortions about the self. *Journal of Abnormal Psychology, 111*, 77-87. http://dx.doi.org/10.1037//021-843x.111.1.77

*McGuire, S., Manke, B., Saudino, K. J., 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. http://dx.doi.org/10.1111/1467-8624.00094

McGuire, S., Neiderhiser, J. M., Reiss, D., Hetherington, E. M., & Plomin, R. (1994). Genetic and environmental influences on perceptions of self-worth and competence in adolescence: A study of twins, full siblings, and step-siblings. *Child Development, 65*, 785-799. http://dx.doi.org/10.1111/j.1467-8624.1994.tb00783.x

*McKenzie, A. A. (2003). The role of primary prevention in immigrant children's development. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 64*(10-A), 3594.

McLeod, B. D., & Weisz, J. R. (2004). Using dissertations to examine potential bias in child and adolescent clinical trials. *Journal of Consulting and Clinical Psychology*, *72*, 235-251. http://dx.doi.org/10.1037/0022-006x.72.2.235

*McLeod, J. D., & Owens, T. J. (2004). Psychological well-being in the early life course: Variations by socioeconomic status, gender, and race/ethnicity. *Social Psychology Quarterly, 67*, 257-278. http://dx.doi.org/10.1177/019027250406700303

McMullin, J. A., & Cairney, J. (2004). Self-esteem and the intersection of age, class, and gender. *Journal of Aging Studies, 18,* 75-90. http://dx.doi.org/10.1016/j.jaging.2003.09.006

*Mellanby, J., Zimdars, A., & Cortina-Borja, M. (2013). Sex differences in degree performance at the University of Oxford. *Learning and Individual Differences*, *26*, 103-111. http://dx.doi.org/10.1016/j.lindif.2013.04.010

Milojev, P., & Sibley, C. G. (2017). Normative personality trait development in adulthood: A 6-year cohort-sequential growth model. *Journal of Personality and Social Psychology, 112,* 510-526. http://dx.doi.org/10.1037/pspp0000121

Mirowsky, J., & Kim, J. (2007). Graphing age trajectories: Vector graphs, synthetic and virtual cohort projections, and cross-sectional profiles of depression. *Sociological Methods and Research, 35*, 497-541. http://dx.doi.org/10.1177/0049124106296015

*Missotten, L. C., Luyckx, K., & Seiffge-Krenke, I. (2013). Family climate of adolescents with and without Type 1 diabetes: Longitudinal associations with psychosocial adaptation. *Journal of Child and Family Studies*, *22*, 344-354. http://dx.doi.org/10.1007/s10826-012-9585-1

*Modecki, K. L., Barber, B. L., & Vernon, L. (2013). Mapping developmental precursors of cyber-aggression: Trajectories of risk predict perpetration and victimization. *Journal of Youth and Adolescence, 42*, 651-661. http://dx.doi.org/10.1007/s10964-012-9887-z

*Molloy, L. E., Ram, N., & Gest, S. D. (2011). The storm and stress (or calm) of early adolescent self-concepts: Within- and between-subjects variability. *Developmental Psychology, 47*, 1589-1607. http://dx.doi.org/10.1037/a0025413

Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry, 12*, 177-196. http://dx.doi.org/10.1207/S15327965pli1204_1

*Morin, A. J. S., Maiano, C., Marsh, H. W., Janosz, M., & Nagengast, B. (2011). The longitudinal interplay of adolescents' self-esteem and body image: A conditional autoregressive latent trajectory analysis. *Multivariate Behavioral Research, 46*, 157-201. http://dx.doi.org/10.1080/00273171.2010.546731

Morin, A. J. S., Maiano, C., Marsh, H. W., Nagengast, B., & Janosz, M. (2013). School life and adolescents’ self-esteem trajectories. *Child Development, 84*, 1967-1988. http://dx.doi.org/10.1111/cdev.12089

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7*, 105-125. http://dx.doi.org/10.1037/1082-989X.7.1.105

Mroczek, D. K., & Kolarz, C. M. (1998). The effect of age on positive and negative affect: A developmental perspective on happiness. *Journal of Personality and Social Psychology, 75*, 1333-1349. http://dx.doi.org/10.1037/0022-3514.75.5.1333

Mroczek, D. K., & Spiro, A. (2005). Change in life satisfaction during adulthood: Findings from the Veterans Affairs Normative Aging Study. *Journal of Personality and Social Psychology, 88*, 189-202. http://dx.doi.org/10.1037/0022-3514.88.1.189

*Mullins, E. R. (1997). Changes in young adolescents' self-perceptions across the transition from elementary to middle school. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 58*(12-A), 4561.

*Mund, M., Finn, C., Hagemeyer, B., Zimmermann, J., & Neyer, F. J. (2015). The dynamics of self-esteem in partner relationships. *European Journal of Personality, 29*, 235-249. http://dx.doi.org/10.1002/per.1984

Mund, M., & Neyer, F. J. (2016). Rising high or falling deep? Pathways of self-esteem in a representative German sample. *European Journal of Personality, 30*, 341-357. http://dx.doi.org/10.1002/per.2063

*Muusses, L. D., Finkenauer, C., Kerkhof, P., & Billedo, C. J. (2014). A longitudinal study of the association between compulsive Internet use and wellbeing. *Computers in Human Behavior, 36*, 21-28. http://dx.doi.org/10.1016/j.chb.2014.03.035

Neiss, M. B., Sedikides, C., & Stevenson, J. (2002). Self-esteem: A behavioural genetic perspective. *European Journal of Personality, 16*, 351-367. http://dx.doi.org/10.1002/per.456

Neiss, M. B., Stevenson, J., Sedikides, C., Kumashiro, M., Finkel, E. J., & Rusbult, C. E. (2005). Executive self, self-esteem, and negative affectivity: Relations at the phenotypic and genotypic level. *Journal of Personality and Social Psychology, 89*, 593-606. http://dx.doi.org/10.1037/0022-3514.89.4.593

*Nordin-Bates, S. M., Quested, E., Walker, I. J., & Redding, E. (2012). Climate change in the dance studio: Findings from the UK centres for advanced training. *Sport, Exercise, and Performance Psychology, 1*, 3-16. http://dx.doi.org/10.1037/a0025316

*O'Dea, J. A. (2006). Self-concept, self-esteem and body weight in adolescent females: A three-year longitudinal study. *Journal of Health Psychology, 11*, 599-611. http://dx.doi.org/10.1177/1359105306065020

*O'Kane, D., Shevlin, M., Armour, C., & Adamson, G. (2010). Testing the effect of an educational intervention: An analysis of the longitudinal changes in academically related psychological variables for Step-Up students. *Irish Journal of Psychology, 31*, 1-10. http://dx.doi.org/10.1080/03033910.2010.10446318

*O'Malley, P. M., & Bachman, J. G. (1983). Self-esteem: Change and stability between ages 13 and 23. *Developmental Psychology, 19*, 257-268. http://dx.doi.org/10.1037/0012-1649.19.2.257

*Oates, G. L. S. C. (2004). The color of the undergraduate experience and the Black self-concept: Evidence from longitudinal data. *Social Psychology Quarterly, 67*, 16-32. http://dx.doi.org/10.1177/019027250406700104

*Ohannessian, C. M., Lerner, R. M., von Eye, A., & Lerner, J. V. (1996). Direct and indirect relations between perceived parental acceptance, perceptions of the self, and emotional adjustment during early adolescence. *Family and Consumer Sciences Research Journal, 25*, 159-183. http://dx.doi.org/10.1177/1077727X960252004

*Oppedal, B., Roysamb, E., & Sam, D. L. (2004). The effect of acculturation and social support on change in mental health among young immigrants. *International Journal of Behavioral Development, 24*, 481-494. http://dx.doi.org/10.1080/01650250444000126

Orth, U. (2018). The family environment in early childhood has a long-term effect on self-esteem: A longitudinal study from birth to age 27 years. *Journal of Personality and Social Psychology*, 114, 637-655. http://dx.doi.org/10.1037/pspp0000143

Orth, U. (2017b). The lifespan development of self-esteem. In J. Specht (Ed.), *Personality development across the lifespan* (pp. 181-195). London, UK: Elsevier. http://dx.doi.org/10.1016/B978-0-12-804674-6.00012-0

Orth, U., Erol, R. Y., Ledermann, T., & Grob, A. (2018). Codevelopment of well-being and self-esteem in romantic partners: Disentangling the effects of mutual influence and shared environment. *Developmental Psychology, 54*, 151-166. http://dx.doi.org/10.1037/dev0000400

*Orth, U., & Luciano, E. C. (2015). Self-esteem, narcissism, and stressful life events: Testing for selection and socialization. *Journal of Personality and Social Psychology*, *109*, 707-721. http://dx.doi.org/10.1037/pspp0000049

*Orth, U., Maes, J., & Schmitt, M. (2015). Self-esteem development across the life span: A longitudinal study with a large sample from Germany. *Developmental Psychology, 51*, 248-259. http://dx.doi.org/10.1037/a0038481

Orth, U., & Robins, R. W. (2013). Understanding the link between low self-esteem and depression. *Current Directions in Psychological Science, 22*, 455-460. http://dx.doi.org/10.1177/0963721413492763

Orth, U., & Robins, R. W. (2014). The development of self-esteem. *Current Directions in Psychological Science, 23*, 381-387. http://dx.doi.org/10.1177/0963721414547414

Orth, U., & Robins, R. W. (in press). Development of self-esteem across the lifespan. In D. P. McAdams, R. L. Shiner, & J. L. Tackett (Eds.), *Handbook of personality development*. New York, NY: Guilford.

Orth, U., Robins, R. W., Meier, L. L., & Conger, R. D. (2016). Refining the vulnerability model of low self-esteem and depression: Disentangling the effects of genuine self-esteem and narcissism. *Journal of Personality and Social Psychology, 110*, 133-149. http://dx.doi.org/10.1037/pspp0000038

*Orth, U., Robins, R. W., & Roberts, B. W. (2008). Low self-esteem prospectively predicts depression in adolescence and young adulthood. *Journal of Personality and Social Psychology, 95*, 695-708. http://dx.doi.org/10.1037/0022-3514.95.3.695

Orth, U., Robins, R. W., Trzesniewski, K. H., Maes, J., & Schmitt, M. (2009). Low self-esteem is a risk factor for depressive symptoms from young adulthood to old age. *Journal of Abnormal Psychology, 118*, 472-478. http://dx.doi.org/10.1037/a0015922

*Orth, U., Robins, R. W., & Widaman, K. F. (2012). Life-span development of self-esteem and its effects on important life outcomes. *Journal of Personality and Social Psychology, 102*, 1271-1288. http://dx.doi.org/10.1037/a0025558

*Orth, U., Robins, R. W., Widaman, K. F., & Conger, R. D. (2014). Is low self-esteem a risk factor for depression? Findings from a longitudinal study of Mexican-origin youth. *Developmental Psychology, 50*, 622-633. http://dx.doi.org/10.1037/a0033817

*Orth, U., Trzesniewski, K. H., & Robins, R. W. (2010). Self-esteem development from young adulthood to old age: A cohort-sequential longitudinal study. *Journal of Personality and Social Psychology, 98*, 645-658. http://dx.doi.org/10.1037/a0018769

*Ostrowsky, M. K. (2006). Extending Khantzian's self-medication hypothesis: An examination of low self-esteem, depression, alcohol use, marijuana use, and violent behavior. *Dissertation Abstracts International Section A: Humanities and Social Sciences, 68*(2-A), 739.

Paulhus, D. L., Robins, R. W., Trzesniewski, K. H., & Tracy, J. L. (2004). Two replicable suppressor situations in personality research. *Multivariate Behavioral Research, 39*, 303-328. http://dx.doi.org/10.1207/s15327906mbr3902_7

Pettit, J. W., & Joiner, T. E. (2001). Negative life events predict negative feedback seeking as a function of impact on self-esteem. *Cognitive Therapy and Research, 25*, 733-741. http://dx.doi.org/10.1023/A:1012919306708

*Pinquart, M., Silbereisen, R. K., & Wiesner, M. (2004). Changes in discrepancies between desired and present states of developmental tasks in adolescence: A 4-process model. *Journal of Youth and Adolescence, 33*, 467-477. http://dx.doi.org/10.1023/B:JOYO.0000048062.81471.5b

*Pomerantz, E. M., & Dong, W. (2006). Effects of mothers' perceptions of children's competence: The moderating role of mothers' theories of competence. *Developmental Psychology, 42*, 950-961. http://dx.doi.org/10.1037/0012-1649.42.5.950

*Pomerantz, E. M., & Rudolph, K. D. (2003). What ensues from emotional distress? Implications for competence estimation. *Child Development, 74*, 329-345. http://dx.doi.org/10.1111/1467-8624.7402001

*Poorthuis, A. M. G., Thomaes, S., van Aken, M. A. G., Denissen, J. J. A., & Orobio de Castro, B. (2014). Dashed hopes, dashed selves? A sociometer perspective on self-esteem change across the transition to secondary school. *Social Development, 23*, 770-783. http://dx.doi.org/10.1111/sode.12075

*Prawat, R. S., Jones, H., & Hampton, J. (1979). Longitudinal study of attitude development in pre-, early, and later adolescent samples. *Journal of Educational Psychology, 71*, 363-369. http://dx.doi.org/10.1037/0022-0663.71.3.363

Preacher, K. J., Wichman, A. L., MacCallum, R. C., & Briggs, N. E. (2008). *Latent growth curve modeling*. Los Angeles: Sage.

*Puckett, M. B. (2009). Rejection sensitivity and interpersonal relationship difficulties: Depression, loneliness, and self-esteem as mediating factors. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 70*(7-B), 4538.

Pullmann, H., Allik, J., & Realo, A. (2009). Global self-esteem across the life span: A cross-sectional comparison between representative and self-selected internet samples. *Experimental Aging Research, 35*, 20-44. http://dx.doi.org/10.1080/03610730802544708

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org.

*Racicot, S., McGrath, J. J., Karp, I., & O'Loughlin, J. (2013). Predictors of nicotine dependence symptoms among never-smoking adolescents: A longitudinal analysis from the Nicotine Dependence in Teens Study. *Drug and Alcohol Dependence, 130*, 38-44. http://dx.doi.org/10.1016/j.drugalcdep.2012.10.006

*Radin, S. M. (2005). The development of drinking in urban American Indian adolescents: A longitudinal examination of self-derogation theory. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 66*(12-B), 6933.

*Raevuori, A., Dick, D. M., Keski-Rahkonen, A., Pulkkinen, L., Rose, R. J., Rissanen, A., . . . Silventoinen, K. (2007). Genetic and environmental factors affecting self-esteem from age 14 to 17: A longitudinal study of Finnish twins. *Psychological Medicine, 37*, 1625-1633. http://dx.doi.org/10.1017/S0033291707000840

*Räikkönen, E., Kokko, K., & Rantanen, J. (2011). Timing of adult transitions: Antecedents and implications for psychological functioning. *European Psychologist, 16*, 314-323. http://dx.doi.org/10.1027/1016-9040/a000050

Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 295-315). New York, NY: Russell Sage Foundation.

*Raustorp, A., Archer, T., Svensson, K., Perlinger, T., & Alricsson, M. (2009). Physical self-esteem, a five year follow-up study on Swedish adolescents. *International Journal of Adolescent Medicine and Health, 21*, 497-507. http://dx.doi.org/10.1515/IJAMH.2009.21.4.497

*Reddy, R., Rhodes, J. E., & Mulhall, P. (2003). The influence of teacher support on student adjustment in the middle school years: A latent growth curve study. *Development and Psychopathology, 15*, 119-138. http://dx.doi.org/10.1017.S0954579403000075

*Reinert, D. F. (2005). Spirituality, self-representations, and attachment to parents: A longitudinal study of Roman Catholic college seminarians. *Counseling and Values, 49*, 226-238. http://dx.doi.org/10.1002/j.2161-007X.2005.tb01025.x

*Reitz, A. K., Motti-Stefanidi, F., & Asendorpf, J. B. (2016). Me, us, and them: Testing sociometer theory in a socially diverse real-life context. *Journal of Personality and Social Psychology, 110*, 908-920. http://dx.doi.org/10.1037/pspp0000073

*Reitzes, D. C., & Mutran, E. J. (2006). Self and health: Factors that encourage self-esteem and functional health. *Journal of Gerontology: Social Sciences, 61B*, S44-S51. http://dx.doi.org/10.1093/geronb/61.1.S44

*Repetto, P. B., Caldwell, C. H., & Zimmerman, M. A. (2004). Trajectories of depressive symptoms among high risk African-American adolescents. *Journal of Adolescent Health, 35*, 468-477. http://dx.doi.org/10.1016/j.jadohealth.2003.12.007

*Reynolds, B. M. (2009). Pubertal timing and internalizing symptoms across the middle school years: Peer context influences in measurement and the search for mechanisms. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 71*(4-B), 2726.

*Reynolds, B. M., & Juvonen, J. (2012). Pubertal timing fluctuations across middle school: implications for girls' psychological health. *Journal of Youth and Adolescence, 41*, 677-690. http://dx.doi.org/10.1007/s10964-011-9687-x

*Rhodes, J. E., Reddy, R., & Grossman, J. B. (2005). The protective influence of mentoring on adolescents' substance use: Direct and indirect pathways. *Applied Developmental Science, 9*, 31-47. http://dx.doi.org/10.1207/s1532480xads0901_4

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*, 1-25. http://dx.doi.org/10.1037/0033-2909.132.1.1

Roberts, B. W., & Wood, D. (2006). Personality development in the context of the neo-socioanalytic model of personality. In D. K. Mroczek & T. D. Little (Eds.), *Handbook of personality development* (pp. 11-39). Mahwah, NJ: Erlbaum.

Roberts, B. W., Wood, D., & Caspi, A. (2008). The development of personality traits in adulthood. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 375-398). New York, NY: Guilford.

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. http://dx.doi.org/10.1177/0146167201272002

Robins, R. W., Tracy, J. L., Trzesniewski, K., Potter, J., & Gosling, S. D. (2001). Personality correlates of self-esteem. *Journal of Research in Personality, 35*, 463-482. http://dx.doi.org/10.1006/jrpe.2001.2324

Robins, R. W., & Trzesniewski, K. H. (2005). Self-esteem development across the lifespan. *Current Directions in Psychological Science, 14*, 158-162. http://dx.doi.org/10.1111/j.0963-7214.2005.00353.x

Robins, R. W., Trzesniewski, K. H., Tracy, J. L., Gosling, S. D., & Potter, J. (2002). Global self-esteem across the life span. *Psychology and Aging, 17*, 423-434. http://dx.doi.org/10.1037/0882-7974.17.3.423

*Rönnau-Böse, M., & Fröhlich-Gildhoff, K. (2009). The promotion of resilience: A person-centered perspective of prevention in early childhood institutions. *Person-Centered and Experiential Psychotherapies*, *8*, 299-318. http://dx.doi.org/10.1080/14779757.2009.9688495

*Rosario, M., Schrimshaw, E. W., & Hunter, J. (2011). Different patterns of sexual identity development over time: Implications for the psychological adjustment of lesbian, gay, and bisexual youths. *Journal of Sex Research, 48*, 3-15. http://dx.doi.org/10.1080/00224490903331067

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

*Rosenfeld, L. B. (2004). A novel approach to the study of aggression: The importance of target specificity in behavior and social cognitions. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 65*(11-B), 6073.

*Rotenberg, K. J., McDougall, P., Boulton, M. J., Vaillancourt, T., Fox, C., & Hymel, S. (2004). Cross-sectional and longitudinal relations among peer-reported trustworthiness, social relationships, and psychological adjustment in children and early adolescents from the United Kingdom and Canada. *Journal of Experimental Child Psychology, 88*, 46-67. http://dx.doi.org/10.1016/j.jecp.2004.01.005

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. http://dx.doi.org/10.1037/0012-1649.16.2.105

*Rueger, S. Y., Malecki, C. K., & Demaray, M. K. (2010). Relationship between multiple sources of perceived social support and psychological and academic adjustment in early adolescence: Comparisons across gender. *Journal of Youth and Adolescence, 39*, 47-61. http://dx.doi.org/10.1007/s10964-008-9368-6

*Sallquist, J., Eisenberg, N., French, D. C., Purwono, U., & Suryanti, T. A. (2010). Indonesian adolescents' spiritual and religious experiences and their longitudinal relations with socioemotional functioning. *Developmental Psychology, 46*, 699-716. http://dx.doi.org/10.1037/a0018879

*Schindler, H. S. (2010). The importance of parenting and financial contributions in promoting fathers' psychological health. *Journal of Marriage and Family, 72*, 318-332. http://dx.doi.org/10.1111/j.1741-3737.2010.00702.x

*Schneider, M., Dunton, G. F., & Cooper, D. M. (2008). Physical activity and physical self-concept among sedentary adolescent females: An intervention study. *Psychology of Sport and Exercise, 9,* 1-14. http://dx.doi.org/10.1016/j.psychsport.2007.01.003

*Seidman, E., Allen, L., Aber, J. L., Mitchell, C., & Feinman, J. (1994). The impact of school transitions in early adolescence on the self-system and perceived social context of poor urban youth. *Child Development, 65,* 507-522. http://dx.doi.org/10.1111/j.1467-8624.1994.tb00766.x

*Seiffge-Krenke, I., & Gelhaar, T. (2008). Does successful attainment of developmental tasks lead to happiness and success in later developmental tasks? A test of Havighurst's (1948) theses. *Journal of Adolescence, 31*, 33-52. http://dx.doi.org/10.1016/j.adolescence.2007.04.002

*Settles, I. H., Jellison, W. A., & Pratt-Hyatt, J. S. (2009). Identification with multiple social groups: The moderating role of identity change over time among women-scientists. *Journal of Research in Personality, 43*, 856-867. http://dx.doi.org/10.1016/j.jrp.2009.04.005

*Shapka, J. D., & Keating, D. P. (2005). Structure and change in self-concept during adolescence. *Canadian Journal of Behavioural Science, 37*, 83-96. http://dx.doi.org/10.1037/h0087247

Shaw, B. A., Liang, J., & Krause, N. (2010). Age and race differences in the trajectory of self-esteem. *Psychology and Aging*, *25*, 84-94. http://dx.doi.org/10.1037/a0018242

*Shek, D. T. (2007). A longitudinal study of perceived parental psychological control and psychological well-being in Chinese adolescents in Hong Kong. *Journal of Clinical Psychology, 63*, 1-22. http://dx.doi.org/10.1002/jclp.20331

*Sher-Censor, E., Parke, R. D., & Coltrane, S. (2011). Parents' promotion of psychological autonomy, psychological control, and Mexican-American adolescents' adjustment. *Journal of Youth and Adolescence, 40*, 620-632. http://dx.doi.org/10.1007/s10964-010-9552-3

*Shim, S. S., Ryan, A. M., & Cassady, J. (2012). Changes in self-esteem across the first year in college: The role of achievement goals. *Educational Psychology, 32*, 149-167. http://dx.doi.org/10.1080/01443410.2011.627837

*Shoshani, A., & Steinmetz, S. (2014). Positive psychology at school: A school-based intervention to promote adolescents' mental health and well-being. *Journal of Happiness Studies, 15*, 1289-1311. http://dx.doi.org/10.1007/s10902-013-9476-1

*Siffert, A., Schwarz, B., & Stutz, M. (2012). Marital conflict and early adolescents' self-evaluation: The role of parenting quality and early adolescents' appraisals. *Journal of Youth and Adolescence*, *41*, 749-763. http://dx.doi.org/10.1007/s10964-011-9703-1

Simmons, R. G., Blyth, D. A., Van Cleave, E. F., & Bush, D. M. (1979). Entry into early adolescence: The impact of school structure, puberty, and early dating on self-esteem. *American Sociological Review, 44*, 948-967. http://dx.doi.org/10.2307/2094719

Simmons, R. G., & Rosenberg, F. (1975). Sex, sex roles, and self-image. *Journal of Youth and Adolescence, 4*, 229-258. http://dx.doi.org/10.1007/BF01537165

*Simpkins, S. D., Eccles, J. S., & Becnel, J. N. (2008). The mediational role of adolescents' friends in relations between activity breadth and adjustment. *Developmental Psychology, 44*, 1081-1094. http://dx.doi.org/10.1037/0012-1649.44.4.1081

*Slonim-Nevo, V., Mirsky, J., Rubinstein, L., & Nauck, B. (2009). The impact of familial and environmental factors on the adjustment of immigrants: A longitudinal study. *Journal of Family Issues, 30*, 92-123. http://dx.doi.org/10.1177/0192513x08324575

*Slutzky, C. B., & Simpkins, S. D. (2009). The link between children's sport participation and self-esteem: Exploring the mediating role of sport self-concept. *Psychology of Sport and Exercise, 10*, 381-389. http://dx.doi.org/10.1016/j.psychsport.2008.09.006

*Smetana, J. G., Campione-Barr, N., & Daddis, C. (2004). Longitudinal development of family decision making: Defining healthy behavioral autonomy for middle-class African American adolescents. *Child Development, 75*, 1418-1434. http://dx.doi.org/10.1111/j.1467-8624.2004.00749.x

Soto, C. J. (2016). The Little Six personality dimensions from early childhood to early adulthood: Mean-level age and gender differences in parents' reports. *Journal of Personality, 84*, 409-422. http://dx.doi.org/10.1111/jopy.12168

Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology, 100*, 330-348. http://dx.doi.org/10.1037/a0021717

Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin*, *139*, 213-240. http://dx.doi.org/10.1037/a0028931

Sowislo, J. F., Orth, U., & Meier, L. L. (2014). What constitutes vulnerable self-esteem? Comparing the prospective effects of low, unstable, and contingent self-esteem on depressive symptoms. *Journal of Abnormal Psychology, 123*, 737-753. http://dx.doi.org/10.1037/a0037770

Specht, J., Bleidorn, W., Denissen, J. J. A., Hennecke, M., Hutteman, R., Kandler, C., . . . Zimmermann, J. (2014). What drives adult personality development? A comparison of theoretical perspectives and empirical evidence. *European Journal of Personality, 28*, 216-230. http://dx.doi.org/10.1002/per.1966

Specht, J., Egloff, B., & Schmukle, S. C. (2011). Stability and change of personality across the life course: The impact of age and major life events on mean-level and rank-order stability of the Big Five. *Journal of Personality and Social Psychology, 101*, 862-882. http://dx.doi.org/10.1037/a0024950

Specht, J., Egloff, B., & Schmukle, S. C. (2013). Everything under control? The effects of age, gender, and education on trajectories of perceived control in a nationally representative German sample. *Developmental Psychology*, *49*, 353-364. http://dx.doi.org/10.1037/a0028243

*Spray, C. M., Warburton, V. E., & Stebbings, J. (2013). Change in physical self-perceptions across the transition to secondary school: Relationships with perceived teacher-emphasised achievement goals in physical education. *Psychology of Sport and Exercise, 14*, 662-669. http://dx.doi.org/10.1016/j.psychsport.2013.05.001

Srivastava, S., & Beer, J. S. (2005). How self-evaluations relate to being liked by others: Integrating sociometer and attachment perspectives. *Journal of Personality and Social Psychology*, *89*, 966-977. http://dx.doi.org/10.1037/0022-3514.89.6.966

Starr, L. R., & Davila, J. (2008). Excessive reassurance seeking, depression, and interpersonal rejection: A meta-analytic review. *Journal of Abnormal Psychology, 117*, 762-775. http://dx.doi.org/10.1037/a0013866

*Steiger, A. E., Allemand, M., Robins, R. W., & Fend, H. A. (2014). Low and decreasing self-esteem during adolescence predict adult depression two decades later. *Journal of Personality and Social Psychology, 106*, 325-338. http://dx.doi.org/10.1037/a0035133

*Steiger, A. E., Fend, H. A., & Allemand, M. (2015). Testing the vulnerability and scar models of self-esteem and depressive symptoms from adolescence to middle adulthood and across generations. *Developmental Psychology, 51*, 236-247. http://dx.doi.org/10.1037/a0038478

*Steinfield, C., Ellison, N. B., & Lampe, C. (2008). Social capital, self-esteem, and use of online social network sites: A longitudinal analysis. *Journal of Applied Developmental Psychology*, *29*, 434-445. http://dx.doi.org/10.1016/j.appdev.2008.07.002

*Steinhausen, H. C., Gavez, S., & Winkler Metzke, C. (2005). Psychosocial correlates, outcome, and stability of abnormal adolescent eating behavior in community samples of young people. *International Journal of Eating Disorders, 37*, 119-126. http://dx.doi.org/10.1002/eat.20077

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. http://dx.doi.org/10.1037/0022-0663.76.1.75

Sutton, A. J. (2009). Publication bias. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 435-452). New York, NY: Russell Sage Foundation.

Swann, W. B., Chang-Schneider, C., & McClarty, K. L. (2007). Do people's self-views matter? *American Psychologist*, *62*, 84-94. http://dx.doi.org/10.1037/0003-066X.62.2.84

Swann, W. B., Chang-Schneider, C., & McClarty, K. L. (2008). Yes, cavalier attitudes can have pernicious consequences. *American Psychologist*, *63*, 65-66. http://dx.doi.org/10.1037/0003-066x.63.1.65

*Syed, M., & Azmitia, M. (2009). Longitudinal trajectories of ethnic identity during the college years. *Journal of Research on Adolescence, 19*, 601-624. http://dx.doi.org/10.1111/j.1532-7795.2009.00609.x

Terracciano, A., McCrae, R. R., Brant, L. J., & Costa, P. T. (2005). Hierarchical linear modeling analyses of the NEO-PI-R scales in the Baltimore Longitudinal Study of Aging. *Psychology and Aging, 20*, 493-506. http://dx.doi.org/10.1037/0882-7974.20.3.493

Tetzner, J., Becker, M., & Baumert, J. (2016). Still doing fine? The interplay of negative life events and self-esteem during young adulthood. *European Journal of Personality, 30*, 358-373. http://dx.doi.org/10.1002/per.2066

*Tiggemann, M. (2005). Body dissatisfaction and adolescent self-esteem: Prospective findings. *Body Image*, *2*, 129-135. http://dx.doi.org/10.1016/j.bodyim.2005.03.006

Tracy, J. L., Cheng, J. T., Robins, R. W., & Trzesniewski, K. H. (2009). Authentic and hubristic pride: The affective core of self-esteem and narcissism. *Self and Identity*, *8*, 196-213. http://dx.doi.org/10.1080/15298860802505053

Trzesniewski, K. H., & Donnellan, M. B. (2010). Rethinking “Generation Me”: A study of cohort effects from 1976-2006. *Perspectives on Psychological Science*, *5*, 58-75. http://dx.doi.org/10.1177/1745691609356789

*Trzesniewski, K. H., Donnellan, M. B., Moffitt, T. E., Robins, R. W., Poulton, R., & Caspi, A. (2006). Low self-esteem during adolescence predicts poor health, criminal behavior, and limited economic prospects during adulthood. *Developmental Psychology, 42*, 381-390. http://dx.doi.org/10.1037/0012-1649.42.2.381

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2003). Stability of self-esteem across the life span. *Journal of Personality and Social Psychology, 84*, 205-220. http://dx.doi.org/10.1037/0022-3514.84.1.205

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2008). Is "Generation Me" really more narcissistic than previous generations? *Journal of Personality, 76*, 903-917. http://dx.doi.org/10.1111/j.1467-6494.2008.00508.x

Trzesniewski, K. H., Donnellan, M. B., & Robins, R. W. (2013). Development of self-esteem. In V. Zeigler-Hill (Ed.), *Self-esteem* (pp. 60-79). London, UK: Psychology Press.

*Turner, H. A., Finkelhor, D., & Ormrod, R. (2010). The effects of adolescent victimization on self-concept and depressive symptoms. *Child Maltreatment, 15*, 76-90. http://dx.doi.org/10.1177/1077559509349444

Twenge, J. M. (2006). *Generation Me: Why today's young Americans are more confident, assertive, entitled – and more miserable than ever before*. New York, NY: Free Press.

Twenge, J. M., & Campbell, W. K. (2001). Age and birth cohort differences in self-esteem: A cross-temporal meta-analysis. *Personality and Social Psychology Review, 5*, 321-344. http://dx.doi.org/10.1207/S15327957PSPR0504_3

Twenge, J. M., & Campbell, W. K. (2010). Birth cohort differences in the Monitoring the Future Dataset and elsewhere: Further evidence for Generation Me—Commentary on Trzesniewski & Donnellan (2010). *Perspectives on Psychological Science, 5*, 81-88. http://dx.doi.org/10.1177/1745691609357015

Twenge, J. M., Carter, N. T., & Campbell, W. K. (2017). Age, time period, and birth cohort differences in self-esteem: Reexamining a cohort-sequential longitudinal study. *Journal of Personality and Social Psychology, 112,* e9-e17. http://dx.doi.org/10.1037/pspp0000122

Twenge, J. M., Konrath, S., Foster, J. D., Campbell, W. K., & Bushman, B. J. (2008). Egos inflating over time: A cross-temporal meta-analysis of the Narcissistic Personality Inventory. *Journal of Personality, 76,* 875-901. http://dx.doi.org/10.1111/j.1467-6494.2008.00507.x

*Udell, W., Sandfort, T., Reitz, E., Bos, H., & Dekovic, M. (2010). The relationship between early sexual debut and psychosocial outcomes: A longitudinal study of Dutch adolescents. *Archives of Sexual Behavior, 39,* 1133-1145. http://dx.doi.org/10.1007/s10508-009-9590-7

*Umaña-Taylor, A. J., Gonzales-Backen, M. A., & Guimond, A. B. (2009). Latino adolescents' ethnic identity: Is there a developmental progression and does growth in ethnic identity predict growth in self-esteem? *Child Development, 80,* 391-405. http://dx.doi.org/10.1111/j.1467-8624.2009.01267.x

*Updegraff, K. A., Perez-Brena, N. J., Umaña-Taylor, A. J., Jahromi, L. B., & Harvey-Mendoza, E. C. (2013). Mothers' trajectories of depressive symptoms across Mexican-origin adolescent daughters' transition to parenthood. *Journal of Family Psychology, 27*, 376-386. http://dx.doi.org/10.1037/a0032909

*van Tuijl, L. A., de Jong, P. J., Sportel, B. E., de Hullu, E., & Nauta, M. H. (2014). Implicit and explicit self-esteem and their reciprocal relationship with symptoms of depression and social anxiety: A longitudinal study in adolescents. *Journal of Behavior Therapy and Experimental Psychiatry, 45*, 113-121. http://dx.doi.org/10.1016/j.jbtep.2013.09.007

*Vanhalst, J., Luyckx, K., Scholte, R. H., Engels, R. C., & Goossens, L. (2013). Low self-esteem as a risk factor for loneliness in adolescence: Perceived - but not actual - social acceptance as an underlying mechanism. *Journal of Abnormal Child Psychology, 41*, 1067-1081. http://dx.doi.org/10.1007/s10802-013-9751-y

*Vasalampi, K., Salmela-Aro, K., & Nurmi, J. E. (2010). Education-related goal appraisals and self-esteem during the transition to secondary education: A longitudinal study. *International Journal of Behavioral Development, 34*, 481-490. http://dx.doi.org/10.1177/0165025409359888

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*, 1-48. http://dx.doi.org/10.18637/jss.v036.i03

Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods, 1*, 112-125. http://dx.doi.org/10.1002/jrsm.11

von Soest, T., Wagner, J., Hansen, T., & Gerstorf, D. (2017). Self-esteem across the second half of life: The role of socioeconomic status, physical health, social relationships, and personality factors. *Journal of Personality and Social Psychology*, Advance online publication. http://dx.doi.org/10.1037/pspp0000123

von Soest, T., Wichstrom, L., & Kvalem, I. L. (2016). The development of global and domain-specific self-esteem from age 13 to 31. *Journal of Personality and Social Psychology, 110*, 592-608. http://dx.doi.org/10.1037/pspp0000060

Wagner, J., Becker, M., Lüdtke, O., & Trautwein, U. (2015). The first partnership experience and personality development: A propensity score matching study in young adulthood.

Social Psychological and Personality Science, 6, 455-463. http://dx.doi.org/10.1177/1948550614566092

Wagner, J., Gerstorf, D., Hoppmann, C., & Luszcz, M. A. (2013). The nature and correlates of self-esteem trajectories in late life. *Journal of Personality and Social Psychology, 105,* 139-153. http://dx.doi.org/10.1037/a0032279

Wagner, J., Hoppmann, C., Ram, N., & Gerstorf, D. (2015). Self-esteem is relatively stable late in life: The role of resources in the health, self-regulation, and social domains. *Developmental Psychology, 51*, 136-149. http://dx.doi.org/10.1037/a0038338

Wagner, J., Lang, F. R., Neyer, F. J., & Wagner, G. G. (2014). Self-esteem across adulthood: The role of resources. *European Journal of Ageing, 11*, 109-119. http://dx.doi.org/10.1007/s10433-013-0299-z

*Wagner, J., Lüdtke, O., Jonkmann, K., & Trautwein, U. (2013). Cherish yourself: Longitudinal patterns and conditions of self-esteem change in the transition to young adulthood. *Journal of Personality and Social Psychology, 104,* 148-163. http://dx.doi.org/10.1037/a0029680

Wagner, J., Lüdtke, O., & Trautwein, U. (2016). Self-esteem is mostly stable across young adulthood: Evidence from latent STARTS models. *Journal of Personality, 84,* 523-535. http://dx.doi.org/10.1111/jopy.12178

Wallace, J., & O'Hara, M. W. (1992). Increases in depressive symptomatology in the rural elderly: Results from a cross-sectional and longitudinal study. *Journal of Abnormal Psychology, 101*, 398-404. http://dx.doi.org/10.1037/0021-843X.101.3.398

*Wallace, J. R., Cunningham, T. F., & Del Monte, V. (1984). Change and stability in self-esteem between late childhood and early adolescence. *Journal of Early Adolescence, 4*, 253-257. http://dx.doi.org/10.1177/0272431684043009

Wängqvist, M., Lamb, M. E., Frisén, A., & Hwang, C. P. (2015). Child and adolescent predictors of personality in early adulthood. *Child Development, 86*, 1253-1261. http://dx.doi.org/10.1111/cdev.12362

*Wargo, J. B. (1999). The influence of grade school characteristics and transition expectations on junior high school adjustment. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 60*(8-B), 4257.

Watson, D., Suls, J., & Haig, J. (2002). Global self-esteem in relation to structural models of personality and affectivity. *Journal of Personality and Social Psychology, 83*, 185-197. http://dx.doi.org/10.1037/0022-3514.83.1.185

*Weed, K., Keogh, D., & Borkowski, J. (2006). Stability of resilience in children of adolescent mothers. *Journal of Applied Developmental Psychology, 27*, 60-77. http://dx.doi.org/10.1016/j.appdev.2005.12.007

*Westaby, J. D., & Lee, B. C. (2003). Antecedents of injury among youth in agricultural settings: A longitudinal examination of safety consciousness, dangerous risk taking, and safety knowledge. *Journal of Safety Research, 34*, 227-240. http://dx.doi.org/10.1016/S0022-4375(03)00030-6

*Whitesell, N. R., Mitchell, C. M., Spicer, P., & The Voices of Indian Teens Project Team. (2009). A longitudinal study of self-esteem, cultural identity, and academic success among American Indian adolescents. *Cultural Diversity and Ethnic Minority Psychology, 15*, 38-50. http://dx.doi.org/10.1037/a0013456

Whiteside-Mansell, L., & Corwyn, R. F. (2003). Mean and covariance structures analyses: An examination of the Rosenberg Self-Esteem Scale among adolescents and adults. *Educational and Psychological Measurement, 63*, 163-173. http://dx.doi.org/10.1177/0013164402239323

Wigfield, A., Eccles, J. S., Mac Iver, D., Reuman, D. A., & Midgley, C. (1991). Transitions during early adolescence: Changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology, 27*, 552-565. http://dx.doi.org/10.1037/0012-1649.27.4.552

*Williams, C., Daley, D., Burnside, E., & Hammond-Rowley, S. (2010). Can trait emotional intelligence and objective measures of emotional ability predict psychopathology across the transition to secondary school? *Personality and Individual Differences, 48*, 161-165. http://dx.doi.org/10.1016/j.paid.2009.09.014

*Wong, D. S., Cheng, C. H., Ngan, R. M., & Ma, S. K. (2011). Program effectiveness of a restorative whole-school approach for tackling school bullying in Hong Kong. *International Journal of Offender Therapy and Comparative Criminology, 55*, 846-862. http://dx.doi.org/10.1177/0306624X10374638

*Wouters, S., Duriez, B., Luyckx, K., Klimstra, T., Colpin, H., Soenens, B., & Verschueren, K. (2013). Depressive symptoms in university freshmen: Longitudinal relations with contingent self-esteem and level of self-esteem. *Journal of Research in Personality, 47*, 356-363. http://dx.doi.org/10.1016/j.jrp.2013.03.001

*Wu, J., Watkins, D., & Hattie, J. (2010). Self-concept clarity: A longitudinal study of Hong Kong adolescents. *Personality and Individual Differences, 48*, 277-282. http://dx.doi.org/10.1016/j.paid.2009.10.011

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

*Yang, H. J. (1997). A prospective longitudinal study of consistency and change in parental influence on children's academic achievement. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 58*(11-B), 6262.

*Yeh, H. C., & Lempers, J. D. (2004). Perceived sibling relationships and adolescent development. *Journal of Youth and Adolescence, 33,* 133-147. http://dx.doi.org/10.1023/B:JOYO.0000013425.86424.0f

*Young, J. F., & Mroczek, D. K. (2003). Predicting intraindividual self-concept trajectories during adolescence. *Journal of Adolescence, 26*, 586-600. http://dx.doi.org/10.1016/S0140-1971(03)00058-7

*Zalta, A. K., & Keel, P. K. (2006). Peer influence on bulimic symptoms in college students. *Journal of Abnormal Psychology, 115*, 185-189. http://dx.doi.org/10.1037/0021-843X.115.1.185

Zeiders, K. H., Umaña-Taylor, A. J., & Derlan, C. L. (2013). Trajectories of depressive symptoms and self-esteem in Latino youths: Examining the role of gender and perceived discrimination. *Developmental Psychology, 49*, 951-963. http://dx.doi.org/10.1037/a0028866

*Zimmerman, M. A., Copeland, L. A., Shope, J. T., & Dielman, T. E. (1997). A longitudinal study of self-esteem: Implications for adolescent development. *Journal of Youth and Adolescence, 26,* 117-141. http://dx.doi.org/10.1023/A:1024596313925

*Zimmermann, F., Schütte, K., Taskinen, P., & Köller, O. (2013). Reciprocal effects between adolescent externalizing problems and measures of achievement. *Journal of Educational Psychology, 105,* 747-761. http://dx.doi.org/10.1037/a0032793

Zuckerman, M., Li, C., & Hall, J. A. (2016). When men and women differ in self-esteem and when they don't: A meta-analysis. *Journal of Research in Personality*, *64*, 34-51. http://dx.doi.org/10.1016/j.jrp.2016.07.007

*Zuckerman, M., & O'Loughlin, R. E. (2009). Narcissism and well-being: A longitudinal perspective. *European Journal of Social Psychology*, *39*, 957-972. http://dx.doi.org/10.1002/ejsp.594

Footnotes

[1] At the time of coding, the qualifications of the coders were as follows: The second author had a Ph.D. in psychology and the third author had a Master's degree in psychology.

[2] Using raw mean change as effect size measure, instead of standardized mean change, was not an option in this meta-analysis. As reported below, the studies used a number of different measures of self-esteem and only standardized effect sizes can be meaningfully compared across measures. Moreover, even within the subset of studies using the Rosenberg scale (i.e., the majority of studies, see below), a standardized effect size measure was needed because many studies did not use the full 10-item Rosenberg scale but an abbreviated version (the most common abbreviated versions included 3, 6, and 7 items). Given that item means differ between items, it is not possible to compare the scale means across different abbreviated versions and the full version of the Rosenberg scale. Moreover, although most studies with the Rosenberg scale used a 4-point response scale, some studies used other response scales such as dichotomous scales and 5-point response scales. Again, it is not possible to compare the scale means across the different response scales used for the Rosenberg scale. For these reasons, it was essential to use a standardized effect size measure for the analyses, even for studies using the Rosenberg scale.

[3] Although effect sizes varied more strongly in childhood and adolescent samples, these samples were not generally smaller than adult samples (see Supplemental Figure S2).

[4] As reported earlier, mean-level change in self-esteem has been examined in a previous meta-analysis (Huang, 2010). To assess whether the available evidence on mean-level change in self-esteem has changed since 2007—the most recent publication year included in Huang's (2010) meta-analytic dataset—we contrasted studies published 2007 and earlier (k = 128) versus

2008 and later (*k* = 203), using mixed-effects meta-regression analysis. The results indicated that effect sizes did not differ between these two sets of studies (estimate = .002, *p* = .874). Thus, the fact that the findings of the present meta-analysis differ from Huang (2010) should not be attributed to differences in the evidence available before and after 2007, but rather to the methodological improvements of the present meta-analysis compared to Huang (2010) as discussed in the Introduction.

[5] Nevertheless, we also tested for the effect of quadratic age, which could capture curvature in the life-span trajectory that deviates from the linear decline of the slope (e.g., the relatively strong decline in old age). Age was centered for these analyses. However, when we added quadratic age to the analysis, neither linear nor quadratic age were significant.

Table 1
Descriptive Information on the Studies Included in the Meta-Analysis

Study	Sample size	Mean age at Time 1	*SD* of age at Time 1	Year of Time 1	Female	Sample type	Country	Ethnicity	Measure of self-esteem
Adachi & Willoughby (2014)	1,492	9.83	acc.	2007	.51	Community	Canada		Rosenb.
Antonishak (2005)	170	13.36	0.66	1998	.53	Community	USA	Other	Harter
Arunkumar et al. (1999)	475	11.00	acc.	1995	.48	Community	USA	Other	Rosenb.
Au et al. (2010)	741	14.50	acc.	2006		Community	China	Asian	Rosenb.
Bachman & O'Malley (1977)	1,608	16.00	acc.	1966	.00	National	USA	Other	Rosenb.
Baldwin & Hoffmann (2002)	108	11.00	acc.	1990	.51	Community	USA	White	Rosenb.
Bao & Jin (2015), control group	80	14.65	0.64	2009	.54	Community	China	Asian	Other
Becker et al. (2014)	2,451	16.00	acc.	1988	.53	Community	USA	Other	Other
Birkelund et al. (2012)	943	14.30	acc.	1991		Community	Norway	White	Rosenb.
Blatny et al. (2015)	83	39.70	acc.	2001	.58	Community	Czech Rep.	White	Rosenb.
Bosacki (2015)	91	6.17	acc.	2003	.57	Community	Canada		Harter
Brenning et al. (2015)	311	12.04	1.41	2011	.54	Community	Belgium	White	Harter
Brown et al. (1998), Black subsample	472	9.00	acc.	1987	1.00	Community	USA	Black	Harter
Brown et al. (1998), White subsample	558	9.00	acc.	1987	1.00	Community	USA	White	Harter
Brummelman et al. (2015)	565	9.56	0.93	2010	.54	Community	Netherlands	White	Harter
Burke & Harrod (2005)	616	25.00	acc.	1991	.50	Community	USA	White	Rosenb.
Burwell & Shirk (2006)	110	13.62	0.52	2002	.58	Community	USA	Other	Harter
Cairns et al. (1990)	2,562	17.00	acc.	1984	.53	Community	Ireland		Harter
Cantin & Boivin (2004)	142	12.50	0.43	1999	.53	Community	Canada		Harter
Caprara et al. (2013), female subsample	109	16.00	acc.	2000	1.00	Community	Italy	White	Rosenb.
Caprara et al. (2013), male subsample	97	16.00	acc.	2000	.00	Community	Italy	White	Rosenb.
Chen et al. (2004), female subsample	258	12.40	0.67	1998	1.00	Community	China	Asian	Harter
Chen et al. (2004), male subsample	248	12.40	0.67	1998	.00	Community	China	Asian	Harter
Chung et al. (2014)	295	18.00	acc.	1992	.60	College	USA	Other	Rosenb.
Colarossi & Eccles (2003), female subsample	125	17.00	acc.	1995	1.00	Community	USA	White	Harter
Colarossi & Eccles (2003), male subsample	92	17.00	acc.	1995	.00	Community	USA	White	Harter
Crocker et al. (2006)	501	14.50	acc.	1998	1.00	Community	Canada		Harter
Davison et al. (2007)	178	11.38	0.28	2001	1.00	Community	USA	White	Harter
Davison et al. (2008)	163	9.34	0.30	2000	1.00	Community	USA	White	Harter
de Araujo & Lagos (2013)	3,363	19.72	1.80	1980	.50	National	USA	Other	Rosenb.
DeRosier (2004), control group	194	8.60	acc.	2000	.49	Community	USA	Other	Harter
DeRosier (2004), nonidentified group	663	8.60	acc.	2000	.49	Community	USA	Other	Harter
Dohnt & Tiggemann (2006)	97	6.91	1.23	2002	1.00	Community	Australia	White	Harter
Donnellan et al. (2007)	409	23.26	0.47	1999	.58	Community	USA	White	Harter
Dorn et al. (2003), female subsample	52	12.00	1.60	1980	1.00	Community	USA	White	Other

Dorn et al. (2003), male subsample	56	12.70	1.30	1980	.00	Community	USA	White	Other
Duckworth et al. (2010)	138	10.52	0.37	2003	.54	Community	USA	Other	Rosenb.
Eisenberg et al. (2006), female subsample, high school	440	12.70	0.74	1998	1.00	Community	USA	Other	Rosenb.
Eisenberg et al. (2006), female subsample, young adults	946	15.80	0.81	1998	1.00	Community	USA	Other	Rosenb.
Eisenberg et al. (2006), male subsample, high school	366	12.80	0.76	1998	.00	Community	USA	Other	Rosenb.
Eisenberg et al. (2006), male subsample, young adults	764	15.90	0.78	1998	.00	Community	USA	Other	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 16–20	177	19.00	1.02	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 21–25	749	23.26	1.36	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 26–30	1,030	28.06	1.43	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 31–35	913	32.96	1.39	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 36–40	831	37.94	1.44	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 41–45	447	42.98	1.39	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 46–50	369	48.02	1.42	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 51–55	305	52.94	1.35	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 56–60	307	58.13	1.40	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 61–65	327	63.00	1.41	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 66–70	257	67.83	1.47	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, female subsample, age 71–75	145	72.92	1.37	1987	1.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 21–25	523	23.50	1.29	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 26–30	899	28.10	1.41	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 31–35	928	33.02	1.41	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 36–40	825	38.03	1.44	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 41–45	537	42.92	1.40	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 46–50	409	47.95	1.43	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 51–55	359	53.02	1.42	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 56–60	302	58.02	1.45	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 61–65	357	63.07	1.37	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 66–70	290	67.90	1.40	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 71–75	206	72.99	1.34	1987	.00	Community	USA	White	Rosenb.
Erol & Orth (2014), Study 2, male subsample, age 76–80	117	77.57	1.37	1987	.00	Community	USA	White	Rosenb.
Falci (2006)	780	15.00	acc.	1988		Community	USA	White	Rosenb.
Fay (2011)	363	14.94	1.09	2006	.59	Community	USA	Other	Harter
Feldman (1999), Asian/female subsample	110	18.00	acc.	1991	1.00	College	USA	Asian	Rosenb.
Feldman (1999), Asian/male subsample	120	18.00	acc.	1991	.00	College	USA	Asian	Rosenb.
Feldman (1999), Black/female subsample	161	18.00	acc.	1991	1.00	College	USA	Black	Rosenb.
Feldman (1999), Black/male subsample	126	18.00	acc.	1991	.00	College	USA	Black	Rosenb.
Feldman (1999), Latino/female subsample	78	18.00	acc.	1991	1.00	College	USA	Hispanic	Rosenb.
Feldman (1999), Latino/male subsample	93	18.00	acc.	1991	.00	College	USA	Hispanic	Rosenb.
Feldman (1999), White/female subsample	136	18.00	acc.	1991	1.00	College	USA	White	Rosenb.
Feldman (1999), White/male subsample	131	18.00	acc.	1991	.00	College	USA	White	Rosenb.

Fenzel (2000)	116	10.80	acc.	1996	.56	Community	USA	White	Harter
Flynn et al. (2014)	228	14.00	acc.	2009	.41	Community	USA	Other	Harter
Foynes et al. (2015), female subsample	230	19.32	2.00	1997	1.00	Community	USA	Other	Rosenb.
Foynes et al. (2015), male subsample	214	19.32	2.00	1997	.00	Community	USA	Other	Rosenb.
Fredricks & Eccles (2008)	1,482	12.27	acc.	1991	.51	Community	USA	Other	Harter
Frijns & Finkenauer (2009)	278	15.60	0.99	2001	.57	Community	Netherlands	White	Rosenb.
Galambos et al. (2006), female subsample	432	18.00	acc.	1984	1.00	Community	Canada	White	Rosenb.
Galambos et al. (2006), male subsample	488	18.00	acc.	1984	.00	Community	Canada	White	Rosenb.
Galliher et al. (2011), female subsample	68	15.24	0.99	2006	1.00	Community	USA	Native	Rosenb.
Galliher et al. (2011), male subsample	65	15.24	0.99	2006	.00	Community	USA	Native	Rosenb.
Gayman et al. (2011)	1,542	20.00	0.94	1998	.45	Community	USA	Other	Rosenb.
Gest et al. (2005)	400	10.00	acc.	2001	.56	Community	USA	White	Harter
Graham et al. (2014)	356	15.80	1.20	2007	1.00	Community	USA	Other	Other
Green et al. (2012)	1,866	13.86	1.28	2007	.39	Community	Australia	White	Marsh
Greene & Way (2005)	135	14.33	0.76	1996	.53	Community	USA	Other	Rosenb.
Gupta et al. (2013), Chinese sample	368	12.20	0.54	2006	.00	Community	China	Asian	Rosenb.
Gupta et al. (2013), US sample	446	11.37	0.58	2005	.00	Community	USA	Other	Rosenb.
Harris et al. (2015), Study 1	982	12.00	acc.	1979		Community	Germany	White	Rosenb.
Harris et al. (2015), Study 2	451	13.00	acc.	1989		Community	USA	White	Rosenb.
Heaven & Ciarrochi (2008), female subsample	437	12.30	0.49	2003	1.00	Community	Australia		Rosenb.
Heaven & Ciarrochi (2008), male subsample	445	12.30	0.49	2003	.00	Community	Australia		Rosenb.
Hirsch & Rapkin (1987)	159	12.00	acc.	1983	.52	Community	USA	Other	Rosenb.
Hoge et al. (1990)	322	12.00	acc.	1983	.55	Community	USA	White	Rosenb.
Hoglund (1995)	41	10.00	acc.	1987	.53	Community	USA	White	Other
Hope et al. (2013)	324	18.00	0.31	2009	.74	College	USA	Black	Rosenb.
Hubbs-Tait et al. (1994)	44	17.70	1.08	1987	1.00	Community	USA	Other	Other
Hutteman et al. (2015)	876	16.00	0.51	2010	.77	Community	Germany		Other
Impett et al. (2011), Study 1	183	14.00	acc.	1998	1.00	Community	USA	Other	Rosenb.
Impett et al. (2011), Study 2	133	14.00	acc.	2001	1.00	Community	USA	Other	Rosenb.
Ireson & Hallam (2009)	1,687	13.50	acc.	2004		Community	UK	White	Marsh
Jackson et al. (2005), Study 1, female subsample	234	19.00	acc.	1993	1.00	College	Canada	White	Rosenb.
Jackson et al. (2005), Study 1, male subsample	117	19.00	acc.	1993	.00	College	Canada	White	Rosenb.
Jackson et al. (2005), Study 2	938	17.50	acc.	1997	.60	Community	Canada	White	Rosenb.
Johnson (2013)	13,401	15.79	1.91	1995		National	USA	Other	Other
Jou (2013)	1,822	18.00	acc.	2005	.49	Community	Taiwan	Asian	Rosenb.
Kahle et al. (1980)	121	16.00	acc.	1974	.00	Community	USA		Rosenb.
Kakihara et al. (2010)	930	15.28	0.98	2004	.47	Community	Sweden	White	Rosenb.
Kaplan (1975)	3,127	13.00	acc.	1971		Community	USA	Other	Rosenb.
Keel et al. (1997), female subsample	80	11.50	acc.	1993	1.00	Community	USA	White	Rosenb.
Keel et al. (1997), male subsample	85	11.50	acc.	1993	.00	Community	USA	White	Rosenb.

Kim & Cicchetti (2009), nonmaltreated group	109	8.50	1.44	2003	.36	Community	USA	Other	Coopers.
Kipp & Weiss (2015)	174	13.50	2.00	2009	1.00	Community	USA	White	Harter
Kistner et al. (2007)	670	9.43	0.99	2003	.55	Community	USA	Other	Harter
Kiviruusu et al. (2015), female subsample	1,067	15.90	0.30	1983	1.00	Community	Finland	White	Rosenb.
Kiviruusu et al. (2015), male subsample	1,105	15.90	0.30	1983	.00	Community	Finland	White	Rosenb.
Klima & Repetti (2008)	247	9.50	acc.	2003	.47	Community	USA	White	Harter
Kuperminc et al. (2004), Black subsample	96	12.50	acc.	1993	.56	Community	USA	Black	Harter
Kuperminc et al. (2004), Latino subsample	122	12.50	acc.	1993	.51	Community	USA	Hispanic	Harter
Kuperminc et al. (2004), White subsample	230	12.50	acc.	1993	.47	Community	USA	White	Harter
Kuster et al. (2012), age 21–25	173	23.48	1.26	2009	.54	Community	Switzerland	White	Rosenb.
Kuster et al. (2012), age 26–30	205	27.70	1.41	2009	.51	Community	Switzerland	White	Rosenb.
Kuster et al. (2012), age 31–35	77	32.49	1.41	2009	.47	Community	Switzerland	White	Rosenb.
Kuster et al. (2012), age 36–40	36	38.25	1.46	2009	.39	Community	Switzerland	White	Rosenb.
Kuster et al. (2012), age 41–45	45	42.73	1.51	2009	.58	Community	Switzerland	White	Rosenb.
Kuster et al. (2012), age 46–50	38	47.90	1.52	2009	.44	Community	Switzerland	White	Rosenb.
Kuster et al. (2012), age 51–55	35	53.14	1.54	2009	.51	Community	Switzerland	White	Rosenb.
Kuster et al. (2013), Dataset 2, age 27–31	135	29.23	1.35	1989	.55	Community	USA	White	Rosenb.
Kuster et al. (2013), Dataset 2, age 32–36	203	33.92	2.36	1989	.51	Community	USA	White	Rosenb.
Kuster et al. (2013), Dataset 2, age 37–41	203	38.78	1.29	1989	.43	Community	USA	White	Rosenb.
Kuzucu et al. (2014)	404	9.00	acc.	1975	.46	Community	USA	Other	Harter
LaGrange et al. (2008), Cohort 1	187	8.45	0.65	2003	.56	Community	USA	Other	Harter
LaGrange et al. (2008), Cohort 2	169	9.56	0.67	2003	.56	Community	USA	Other	Harter
LaGrange et al. (2008), Cohort 3	171	11.35	0.56	2003	.56	Community	USA	Other	Harter
Larsen et al. (2010)	433	12.25	0.47	2000	.55	Community	Netherlands	White	Rosenb.
Le Bars et al. (2009), Study 2	82	16.70	1.10	2004	.45	Community	France	White	Other
Lee (2012)	3,449	13.78	0.42	2003	.50	Community	Korea	Asian	Other
Leeuwis et al. (2015), female subsample	173	11.20	0.40	2011	1.00	Community	Netherlands	White	Harter
Leeuwis et al. (2015), male subsample	157	11.20	0.40	2011	.00	Community	Netherlands	White	Harter
Lemmens et al. (2011)	543	13.90	1.40	2008	.49	Community	Netherlands	White	Rosenb.
Lin et al. (2012)	152	21.00	acc.	1985	.57	Community	Netherlands	White	Other
Luyckx et al. (2013), Study 2	565	18.66	0.66	2002	.85	College	Belgium	White	Rosenb.
Maldonado et al. (2013)	756	13.76	2.58	1983	.49	Community	USA	White	Other
Marsh et al. (1998), kindergarten	127	5.40	0.40	1994		Community	Australia		Marsh
Marsh et al. (1998), Grade 1	139	6.30	0.40	1994		Community	Australia		Marsh
Marsh et al. (1998), Grade 2	130	7.40	0.50	1994		Community	Australia		Marsh
Marsh et al. (2005), Study 3	3,731	15.00	acc.	2000		Community	Australia		Marsh
Marshall et al. (2014)	778	15.41	0.53	2006	.49	Community	Australia	White	Rosenb.
Martin et al. (2013)	643	12.61	1.77	2010	.55	Community	Australia	Other	Marsh
Martins & Harrison (2012)	396	8.74	1.01	2008	.52	Community	USA	Other	Other
Mayzer (2004)	220	7.00	acc.	1994	.00	Community	USA	White	Harter

McCabe et al. (2005)	412	9.28	acc.	1998	.52	Community	Australia	Other	Marsh
McCarthy & Hoge (1982), Grade 7–8	546	13.50	acc.	1976	.45	Community	USA	Other	Other
McCarthy & Hoge (1982), Grade 9–10	775	15.50	acc.	1976	.45	Community	USA	Other	Other
McCarthy & Hoge (1982), Grade 11–12	478	17.50	acc.	1976	.45	Community	USA	Other	Other
McCarty et al. (2007)	331	12.00	0.41	2001	.47	Community	USA	Other	Harter
McGrath & Repetti (2002)	244	9.50	acc.	1997	.47	Community	USA	White	Harter
McGuire et al. (1999)	496	12.90	2.00	1989	.51	Community	USA	Other	Harter
McKenzie (2003), control group	58	7.00	acc.	1996	.75	Community	Canada	Other	Marsh
McLeod & Owens (2004)	547	10.50	acc.	1986	.51	National	USA	Other	Harter
Mellanby et al. (2013), female subsample	356	18.00	acc.	2000	1.00	College	UK	White	Rosenb.
Mellanby et al. (2013), male subsample	340	18.00	acc.	2000	.00	College	UK	White	Rosenb.
Missotten et al. (2012), control group	119	13.90	1.28	1995	.56	Community	Germany	White	Other
Modecki et al. (2013)	1,364	13.00	0.34	2007	.55	Community	Australia	White	Other
Molloy et al. (2011)	152	9.00	acc.	2001	.46	Community	USA	White	Harter
Morin et al. (2011)	1,001	12.62	0.63	2000	.46	Community	Canada	White	Rosenb.
Mullins (1997), female subsample	131	11.00	acc.	1992	1.00	Community	USA	White	Harter
Mullins (1997), male subsample	120	11.00	acc.	1992	.00	Community	USA	White	Harter
Mund et al. (2015), Study 1	186	26.82	3.01	1995	.66	Community	Germany	White	Marsh
Muusses et al. (2014)	398	30.64	4.57	2005	.50	Community	Netherlands	White	Rosenb.
Nordin-Bates et al. (2012)	326	14.41	2.10	2008	.76	Community	UK	White	Rosenb.
O'Dea (2006)	80	12.80	0.60	2001	1.00	Community	Australia	White	Harter
O'Kane et al. (2010), control group	97	16.50	0.50	2004	.64	Community	Ireland	White	Rosenb.
O'Malley & Bachman (1983), class of 1976, 1st sample	113	18.00	acc.	1976		Community	USA		Rosenb.
O'Malley & Bachman (1983), class of 1976, 2nd sample	129	18.00	acc.	1976		Community	USA		Rosenb.
O'Malley & Bachman (1983), class of 1977, 1st sample	127	18.00	acc.	1977		Community	USA		Rosenb.
O'Malley & Bachman (1983), class of 1977, 2nd sample	130	18.00	acc.	1977		Community	USA		Rosenb.
O'Malley & Bachman (1983), class of 1978, 1st sample	132	18.00	acc.	1978		Community	USA		Rosenb.
O'Malley & Bachman (1983), class of 1978, 2nd sample	142	18.00	acc.	1978		Community	USA		Rosenb.
O'Malley & Bachman (1983), class of 1979	148	18.00	acc.	1979		Community	USA		Rosenb.
Oates (2004)	607	18.00	acc.	1971	.00	Community	USA	Black	Rosenb.
Ohannessian et al. (1996), female subsample	103	12.20	0.68	1990	1.00	Community	USA	White	Harter
Ohannessian et al. (1996), male subsample	101	12.20	0.68	1990	.00	Community	USA	White	Harter
Oppedal et al. (2004), female subsample	63	13.00	acc.	2000	1.00	Community	Norway	Other	Rosenb.
Oppedal et al. (2004), male subsample	74	13.00	acc.	2000	.00	Community	Norway	Other	Rosenb.
Orth & Luciano (2015), Study 1	328	21.17	1.93	2010	.50	Community	Switzerland	White	Rosenb.
Orth & Luciano (2015), Study 2, age 18–22	85	20.61	1.31	2011	.61	Community	Switzerland	White	Rosenb.
Orth & Luciano (2015), Study 2, age 23–27	121	24.94	1.34	2011	.56	Community	Switzerland	White	Rosenb.
Orth & Luciano (2015), Study 2, age 28–32	71	29.70	1.53	2011	.37	Community	Switzerland	White	Rosenb.
Orth & Luciano (2015), Study 2, age 33–37	37	34.92	1.36	2011	.46	Community	Switzerland	White	Rosenb.
Orth et al. (2008), Study 1	2,094	15.50	0.50	1994	.50	National	USA	Other	Rosenb.

Orth et al. (2010), age 25–29	317	27.10	1.23	1986	.55	National	USA	Other	Rosenb.
Orth et al. (2010), age 30–34	376	31.91	1.42	1986	.56	National	USA	Other	Rosenb.
Orth et al. (2010), age 35–39	336	36.80	1.42	1986	.60	National	USA	Other	Rosenb.
Orth et al. (2010), age 40–44	268	41.92	1.43	1986	.62	National	USA	Other	Rosenb.
Orth et al. (2010), age 45–49	212	46.89	1.36	1986	.56	National	USA	Other	Rosenb.
Orth et al. (2010), age 50–54	176	51.90	1.42	1986	.57	National	USA	Other	Rosenb.
Orth et al. (2010), age 55–59	221	57.19	1.45	1986	.66	National	USA	Other	Rosenb.
Orth et al. (2010), age 60–64	436	62.03	1.39	1986	.61	National	USA	Other	Rosenb.
Orth et al. (2010), age 65–69	438	66.94	1.46	1986	.65	National	USA	Other	Rosenb.
Orth et al. (2010), age 70–74	311	71.88	1.43	1986	.72	National	USA	Other	Rosenb.
Orth et al. (2010), age 75–79	265	76.75	1.49	1986	.69	National	USA	Other	Rosenb.
Orth et al. (2010), age 80–84	134	81.72	1.44	1986	.71	National	USA	Other	Rosenb.
Orth et al. (2010), age 85–89	51	86.63	1.41	1986	.59	National	USA	Other	Rosenb.
Orth et al. (2012), age 11–15	82	16.79	0.82	1988	.58	Community	USA	White	Rosenb.
Orth et al. (2012), age 16–20	80	21.00	1.27	1988	.62	Community	USA	White	Rosenb.
Orth et al. (2012), age 31–35	323	33.86	1.00	1988	.60	Community	USA	White	Rosenb.
Orth et al. (2012), age 36–40	292	37.55	1.36	1988	.53	Community	USA	White	Rosenb.
Orth et al. (2012), age 41–45	89	42.01	1.13	1988	.46	Community	USA	White	Rosenb.
Orth et al. (2012), age 51–55	81	53.95	1.11	1988	.84	Community	USA	White	Rosenb.
Orth et al. (2012), age 56–60	213	58.05	1.46	1988	.61	Community	USA	White	Rosenb.
Orth et al. (2012), age 61–65	151	62.76	1.37	1988	.50	Community	USA	White	Rosenb.
Orth et al. (2012), age 66–70	78	67.56	1.34	1988	.42	Community	USA	White	Rosenb.
Orth et al. (2012), age 76–80	71	78.16	1.24	1988	.68	Community	USA	White	Rosenb.
Orth et al. (2012), age 81–85	49	82.54	1.38	1988	.48	Community	USA	White	Rosenb.
Orth et al. (2014)	672	10.40	0.60	2006	.50	Community	USA	Hispanic	Marsh
Orth et al. (2015), age 14–18	51	16.96	1.15	1996	.65	Community	Germany	White	Rosenb.
Orth et al. (2015), age 19–23	103	20.88	1.48	1996	.56	Community	Germany	White	Rosenb.
Orth et al. (2015), age 24–28	204	26.08	1.42	1996	.48	Community	Germany	White	Rosenb.
Orth et al. (2015), age 29–33	227	31.13	1.46	1996	.47	Community	Germany	White	Rosenb.
Orth et al. (2015), age 34–38	214	35.90	1.39	1996	.50	Community	Germany	White	Rosenb.
Orth et al. (2015), age 39–43	230	41.17	1.40	1996	.49	Community	Germany	White	Rosenb.
Orth et al. (2015), age 44–48	236	45.93	1.37	1996	.42	Community	Germany	White	Rosenb.
Orth et al. (2015), age 49–53	233	51.12	1.42	1996	.38	Community	Germany	White	Rosenb.
Orth et al. (2015), age 54–58	294	56.00	1.37	1996	.31	Community	Germany	White	Rosenb.
Orth et al. (2015), age 59–63	266	60.72	1.41	1996	.37	Community	Germany	White	Rosenb.
Orth et al. (2015), age 64–68	199	65.83	1.35	1996	.20	Community	Germany	White	Rosenb.
Orth et al. (2015), age 69–73	153	70.79	1.42	1996	.32	Community	Germany	White	Rosenb.
Orth et al. (2015), age 74–78	66	75.32	1.34	1996	.35	Community	Germany	White	Rosenb.
Ostrowsky (2006), female subsample	696	14.50	acc.	1989	1.00	Community	USA	Other	Rosenb.
Ostrowsky (2006), male subsample	710	14.50	acc.	1989	.00	Community	USA	Other	Rosenb.

Pinquart et al. (2004)	980	13.70	1.60	1983	.50	National	Germany	White	Rosenb.
Pomerantz & Dong (2006)	126	10.03	0.83	1997	.44	Community	USA	White	Harter
Pomerantz & Rudolph (2003), female subsample	466	10.25	acc.	1997	1.00	Community	USA	White	Harter
Pomerantz & Rudolph (2003), male subsample	466	10.25	acc.	1997	.00	Community	USA	White	Harter
Poorthuis et al. (2014)	306	12.20	0.44	2010	.53	Community	Netherlands	White	Harter
Prawat et al. (1979), female subsample, preadolescents	56	10.40	1.30	1976	1.00	Community	USA	White	Coopers.
Prawat et al. (1979), female subsample, early adolescents	73	13.70	1.00	1976	1.00	Community	USA	White	Coopers.
Prawat et al. (1979), female subsample, later adolescents	70	16.40	1.00	1976	1.00	Community	USA	White	Coopers.
Prawat et al. (1979), male subsample, preadolescents	56	10.40	1.30	1976	.00	Community	USA	White	Coopers.
Prawat et al. (1979), male subsample, early adolescents	67	13.70	1.00	1976	.00	Community	USA	White	Coopers.
Prawat et al. (1979), male subsample, later adolescents	60	16.40	1.00	1976	.00	Community	USA	White	Coopers.
Puckett (2009), female subsample	215	13.95	0.36	2005	1.00	Community	USA	Other	Rosenb.
Puckett (2009), male subsample	156	13.95	0.36	2005	.00	Community	USA	Other	Rosenb.
Racicot et al. (2013)	847	12.70	0.50	1999	.50	Community	Canada		Rosenb.
Radin (2005), younger cohort	290	11.67	0.69	1988	.50	Community	USA	Native	Harter
Radin (2005), older cohort	283	13.69	0.72	1990	.50	Community	USA	Native	Harter
Raevuori et al. (2007), female subsample	2,062	14.00	acc.	1997	1.00	Community	Finland	White	Rosenb.
Raevuori et al. (2007), male subsample	2,070	14.00	acc.	1997	.00	Community	Finland	White	Rosenb.
Räikkönen et al. (2011)	282	36.00	0.25	1995	.47	Community	Finland	White	Rosenb.
Raustorp et al. (2009), female subsample	36	12.70	acc.	2000	1.00	Community	Sweden	White	Other
Raustorp et al. (2009), male subsample	41	12.70	acc.	2000	.00	Community	Sweden	White	Other
Reddy et al. (2003), female subsample	1,237	12.00	acc.	1995	1.00	Community	USA	White	Other
Reddy et al. (2003), male subsample	1,227	12.00	acc.	1995	.00	Community	USA	White	Other
Reinert (2005)	75	22.10	4.00	2001	.00	College	USA	Other	Rosenb.
Reitz et al. (2016)	1,057	12.70	0.66	2005	.47	Community	Greece	White	Rosenb.
Reitzes & Mutran (2006)	737	60.54	1.85	1992	.52	Community	USA	White	Rosenb.
Repetto et al. (2004)	579	14.61	0.68	1998	.53	Community	USA	Black	Other
Reynolds (2009), Study 1	1,038	11.60	acc.	1998	1.00	Community	USA	Other	Harter
Reynolds & Juvonen (2012)	1,038	11.60	acc.	2000	1.00	Community	USA	Other	Harter
Rhodes et al. (2005), control group	581	12.33	acc.	2000	.35	Community	USA	Other	Harter
Rönnau-Böse & Fröhlich-Gildhoff (2009), control group	102	4.30	1.00	2006	.47	Community	Germany	White	Other
Rosario et al. (2011)	156	18.30	1.65	1993	.49	Community	USA	Other	Rosenb.
Rosenfeld (2004), female subsample	181	13.70	acc.	2002	1.00	Community	USA	Other	Harter
Rosenfeld (2004), male subsample	159	13.70	acc.	2002	.00	Community	USA	Other	Harter
Rotenberg et al. (2004), Study 1	505	9.75	0.64	2000	.53	Community	UK		Harter
Rueger et al. (2010), female subsample	325	13.50	acc.	2006	1.00	Community	USA	Other	Other
Rueger et al. (2010), male subsample	311	13.50	acc.	2006	.00	Community	USA	Other	Other
Sallquist et al. (2010)	205	13.47	0.69	2004	.55	Community	Indonesia	Asian	Harter
Schindler (2010)	538	36.85	acc.	1997	.00	Community	USA	Other	Rosenb.
Schneider et al. (2008), control group	59	15.02	0.77	2004	1.00	Community	USA		Marsh

Seidman et al. (1994)	580	11.40	0.93	1988	.54	Community	USA	Other	Harter
Seiffge-Krenke & Gelhaar (2008), female subsample	80	13.90	1.40	1996	1.00	National	Germany	White	Other
Seiffge-Krenke & Gelhaar (2008), male subsample	66	13.90	1.40	1996	.00	National	Germany	White	Other
Settles et al. (2009)	128	22.20	2.55	2004	1.00	College	USA	Other	Rosenb.
Shapka & Keating (2005)	518	15.50	acc.	2000	.49	Community	Canada		Harter
Shek (2007), female subsample	1,670	12.65	0.88	2003	1.00	Community	China	Asian	Rosenb.
Shek (2007), male subsample	1,331	12.65	0.88	2003	.00	Community	China	Asian	Rosenb.
Sher-Censor et al. (2011)	134	10.83	0.63	2006	.55	Community	USA	Hispanic	Harter
Shim et al. (2012)	311	18.00	acc.	2001	.54	College	USA	Other	Rosenb.
Shoshani & Steinmetz (2014)	501	13.75	0.66	2010	.51	Community	Israel	White	Rosenb.
Siffert et al. (2012)	176	10.61	0.40	2008	.51	Community	Switzerland	White	Harter
Simpkins et al. (2008)	925	14.55	1.31	1994	.51	Community	USA	White	Harter
Slonim-Nevo et al. (2010), Israeli sample, adolescents	61	15.97	1.20	2001	.51	Community	Israel	White	Rosenb.
Slonim-Nevo et al. (2013), German sample, adolescents	134	14.16	2.00	2001	.57	Community	Germany	White	Rosenb.
Slutzky & Simpkins (2009)	987	9.55	1.31	1989	.51	Community	USA	White	Harter
Smetana et al. (2004), adolescent sample	76	13.43	1.39	1997	.50	Community	USA	Black	Harter
Spray et al. (2013)	491	11.29	0.30	2009	.51	Community	UK	White	Marsh
Steiger et al. (2014)	1,527	12.00	acc.	1979	.51	Community	Germany	White	Rosenb.
Steiger et al. (2015)	1,359	16.71	0.73	1983	.51	Community	Germany	White	Other
Steinfield et al. (2008)	92	20.10	1.36	2006	.74	College	USA	White	Rosenb.
Steinhausen et al. (2005), control group	64	14.80	1.03	1994	.77	Community	Switzerland	White	Rosenb.
Syed & Azmitia (2009)	175	18.02	0.43	2003	.62	College	USA	Other	Rosenb.
Tiggemann (2005)	242	14.00	0.94	2000	1.00	Community	Australia	White	Rosenb.
Trzesniewski et al. (2006)	812	11.00	acc.	1983	.49	Community	New Zeal.	White	Rosenb.
Turner et al. (2010)	523	14.50	acc.	2003	.53	Community	USA	White	Rosenb.
Udell et al. (2010), female subsample	258	13.31	0.51	2002	1.00	Community	Netherlands	White	Harter
Udell et al. (2010), male subsample	212	13.31	0.51	2002	.00	Community	Netherlands	White	Harter
Umaña-Taylor et al. (2009), female subsample	160	15.23	0.77	2003	1.00	Community	USA	Hispanic	Rosenb.
Umaña-Taylor et al. (2009), male subsample	163	15.39	0.74	2003	.00	Community	USA	Hispanic	Rosenb.
Updegraff et al. (2013), adolescent sample	204	16.24	0.99	2007	1.00	Community	USA	Hispanic	Rosenb.
van Tuijl et al. (2014)	1,641	13.14	0.75	2006	.53	Community	Netherlands	White	Rosenb.
Vanhalst et al. (2013), Study 1, female subsample	201	15.22	0.60	2002	1.00	Community	Netherlands	White	Rosenb.
Vanhalst et al. (2013), Study 1, male subsample	227	15.22	0.60	2002	.00	Community	Netherlands	White	Rosenb.
Vanhalst et al. (2013), Study 2, female subsample	331	14.95	0.94	2007	1.00	Community	Belgium	White	Rosenb.
Vanhalst et al. (2013), Study 2, male subsample	195	14.95	0.94	2007	.00	Community	Belgium	White	Rosenb.
Vasalampi et al. (2010)	606	16.00	0.34	2004	.48	Community	Finland	White	Rosenb.
Wagner et al. (2013)	4,471	19.60	0.85	2002	.55	Community	Germany	White	Marsh
Wallace et al. (1984), female subsample	38	9.40	acc.	1973	1.00	Community	USA		Coopers.
Wallace et al. (1984), male subsample	32	9.40	acc.	1973	.00	Community	USA		Coopers.
Wargo (1999), female subsample	75	12.00	acc.	1995	1.00	Community	USA	Other	Harter

Wargo (1999), male subsample	46	12.00	acc.	1995	.00	Community	USA	Other	Harter
Weed et al. (2006)	106	22.10	1.27	1995	1.00	Community	USA	Other	Coopers.
Westaby & Lee (2003)	3,081	16.50	acc.	1999	.33	Community	USA		Rosenb.
Whitesell et al. (2009)	1,611	14.00	acc.	1993	.50	Community	USA	Native	Rosenb.
Williams et al. (2010)	274	10.50	acc.	2006	.51	Community	UK	White	Other
Wong et al. (2011), control group	196	13.00	acc.	2004		Community	China	Asian	Other
Wouters et al. (2013), female subsample	381	18.41	1.43	2009	1.00	College	Belgium	White	Rosenb.
Wouters et al. (2013), male subsample	73	18.41	1.43	2009	.00	College	Belgium	White	Rosenb.
Wu et al. (2010)	1,044	15.00	1.70	2006	.48	Community	China	Asian	Other
Yang (1997)	398	12.42	acc.	1989	.50	Community	USA	Other	Rosenb.
Yeh & Lempers (2004)	374	12.37	0.51	1999	.50	Community	USA	White	Rosenb.
Young & Mroczek (2003)	261	15.50	acc.	1999	.45	College	USA	White	Harter
Zalta & Keel (2006)	114	19.80	0.90	2003	.47	College	USA	Other	Rosenb.
Zimmerman et al. (1997)	1,160	12.00	acc.	1990	.50	Community	USA	White	Coopers.
Zimmermann et al. (2013)	1,045	11.00	acc.	2006	.50	Community	Germany	White	Rosenb.
Zuckerman & O'Loughlin (2009)	176	20.00	acc.	2002	.72	College	USA		Rosenb.

Note. Mean age and standard deviation of age are given in years. As described in the Method section, some studies did not report the standard deviation of age, but other information clearly suggested that the sample was sufficiently homogeneous with regard to age (e.g., all participants were children in the same grade). For these studies, the standard deviation is denoted acceptable ("acc."). The column "Female" shows the proportion of female participants. Rosenb. = Rosenberg; Coopers. = Coopersmith.

Table 2

Estimates of Yearly Mean-Level Change in Self-Esteem for Age Groups From 4 to 94 Years

Age group			Weighted mean		Heterogeneity		
(years)	k	N	effect size (d_{year})	95% CI	Q	τ^2	I^2
4–6	2	229	0.072	[−.058, .202]	0.5	0.000	0.0
6–8	5	515	0.035	[−.123, .193]	12.7*	0.022	68.6
8–10	13	5,002	0.016	[−.120, .152]	270.7*	0.059	95.6
10–11	21	7,952	0.079*	[.024, .134]	107.2*	0.012	81.3
11–12	29	11,284	0.011	[−.036, .058]	157.0*	0.012	82.2
12–13	45	19,482	−0.004	[−.041, .034]	255.2*	0.011	82.8
13–14	50	29,441	0.005	[−.034, .044]	492.2*	0.016	90.1
14–15	45	30,856	−0.002	[−.041, .037]	441.1*	0.014	90.0
15–16	47	31,729	0.049*	[.020, .078]	263.0*	0.007	82.5
16–17	43	23,369	0.056*	[.021, .091]	236.8*	0.009	82.3
17–18	24	14,373	0.059*	[.019, .098]	90.5*	0.006	74.6
18–19	26	25,426	0.062*	[.020, .104]	174.9*	0.008	85.7
19–20	33	11,912	0.088*	[.059, .117]	57.4*	0.002	44.3
20–25	35	21,275	0.044*	[.021, .067]	64.0*	0.002	46.8
25–30	16	8,812	0.036*	[.010, .061]	18.9	0.001	20.8
30–40	27	15,745	0.012	[−.004, .027]	22.6	0.000	0.0
40–50	19	7,212	0.011	[−.014, .035]	19.5	0.000	7.7
50–60	16	4,248	0.002	[−.028, .032]	2.3	0.000	0.0
60–70	17	4,953	0.000	[−.028, .028]	3.7	0.000	0.0
70–80	15	3,551	−0.006	[−.039, .027]	2.3	0.000	0.0
80–90	7	998	−0.009	[−.071, .053]	2.1	0.000	0.0
90–94	2	185	−0.097	[−.242, .047]	0.2	0.000	0.0

Note. Computations were made with random-effects models. k = number of samples; N = total number of participants in the k samples; d_{year} = standardized mean change d per year; CI = confidence interval; Q = statistic used in heterogeneity test; τ^2 = estimated amount of total heterogeneity; I^2 = ratio of total heterogeneity by total variability (given in percent).

* $p < .05$.

Table 3

Intercorrelations Among Effect Size and Sample Characteristics

Variable	1	2	3	4	5	6	7	8
1. d_{year}	—	.11	−.14*	−.01	.00	.08	−.05	.16*
2. Mean age[a]	−.06	—	−.34*	.00	.09	−.15*	−.06	.50*
3. Mean year of birth	.05	−.92*	—	.07	−.18*	−.33*	.01	−.04
4. Female (proportion)	−.04	−.04	.06	—	−.06	.01	.01	−.06
5. Sample type[b]	−.04	.32*	−.36*	.02	—	−.03	.00	.01
6. Country[c]	−.06	.12*	−.28*	.02	.13*	—	−.35*	−.23*
7. Ethnicity[d]	−.01	.20*	−.16*	−.02	−.23*	−.33*	—	.01
8. Measure of self-esteem[e]	.14*	.47*	−.39*	−.07	.13*	−.09	.12*	—

Note. Correlations for samples with mean age from 4 to 94 years (i.e., the full observed range of mean age; k = 311) are presented below the diagonal, and correlations for samples with mean age from 10 to 20 years (k = 204) are presented above the diagonal. Due to missing values, in these analyses the number of samples was reduced from 331 to 311 (for mean age from 4 to 94 years) and from 220 to 204 (for mean age from 10 to 20 years). d_{year} = standardized mean change d per year.

[a] Mean age of sample at the center of the observed time interval.

[b] 1 = nationally representative, 0 = other.

[c] 1 = United States, 0 = other.

[d] 1 = White/European, 0 = other.

[e] 1 = Rosenberg scale, 0 = other.

* $p < .05$.

Table 4

Mixed-Effects Meta-Regression Models for Sample Characteristics Predicting Yearly Mean-Level Change (d_{year}) in Self-Esteem

	Samples with mean age from 4 to 94 years (k = 311)			Samples with mean age from 10 to 20 years (k = 204)		
Moderator	Estimate	*SE*	*p*	Estimate	*SE*	*p*
Mean age[a]	−.0025*	.0010	.018	.0011	.0041	.786
Mean year of birth	−.0013	.0009	.142	−.0014	.0010	.185
Female (proportion)	.0000	.0002	.907	.0001	.0002	.643
Sample type[b]	−.0019	.0287	.946	.0053	.0481	.913
Country[c]	−.0221	.0159	.165	.0088	.0200	.660
Ethnicity[d]	−.0077	.0149	.606	−.0066	.0180	.712
Measure of self-esteem[e]	.0513*	.0149	.001	.0408*	.0192	.033

Note. Regression coefficients are unstandardized. Due to missing values, in these analyses the number of samples was reduced from 331 to 311 (for mean age from 4 to 94 years) and from 220 to 204 (for mean age from 10 to 20 years). For mean age from 4 to 94 years: Q_{model} = 18.0 (df = 7, p = .012); $Q_{residual}$ = 1336.9 (df = 303, p < .001); τ^2 = 0.008; I^2 = 77.3%. For mean age from 10 to 20 years: Q_{model} = 11.4 (df = 7, p = .121); $Q_{residual}$ = 1009.1 (df = 196, p < .001); τ^2 = 0.009; I^2 = 80.6%. d_{year} = standardized mean change d per year; k = number of samples; *SE* = standard error; τ^2 = estimated amount of total heterogeneity; I^2 = ratio of total heterogeneity by total variability.

[a] Mean age of sample at the center of the observed time interval.

[b] 1 = nationally representative, 0 = other.

[c] 1 = United States, 0 = other.

[d] 1 = White/European, 0 = other.

[e] 1 = Rosenberg scale, 0 = other.

* p < .05.

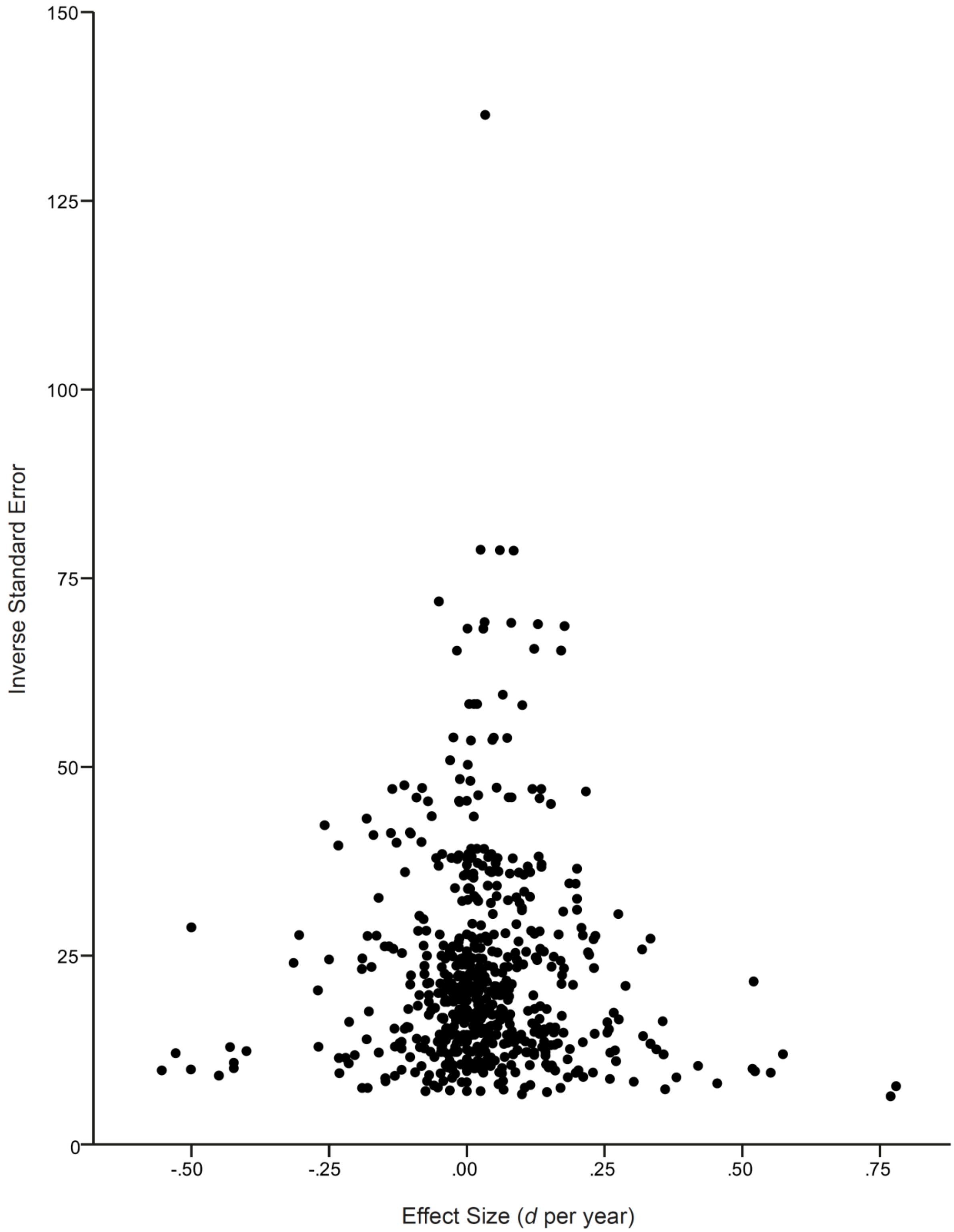

Figure 1. Funnel graph displaying the relation between the inverse standard error and effect size (i.e., standardized mean change *d* per year, d_{year}).

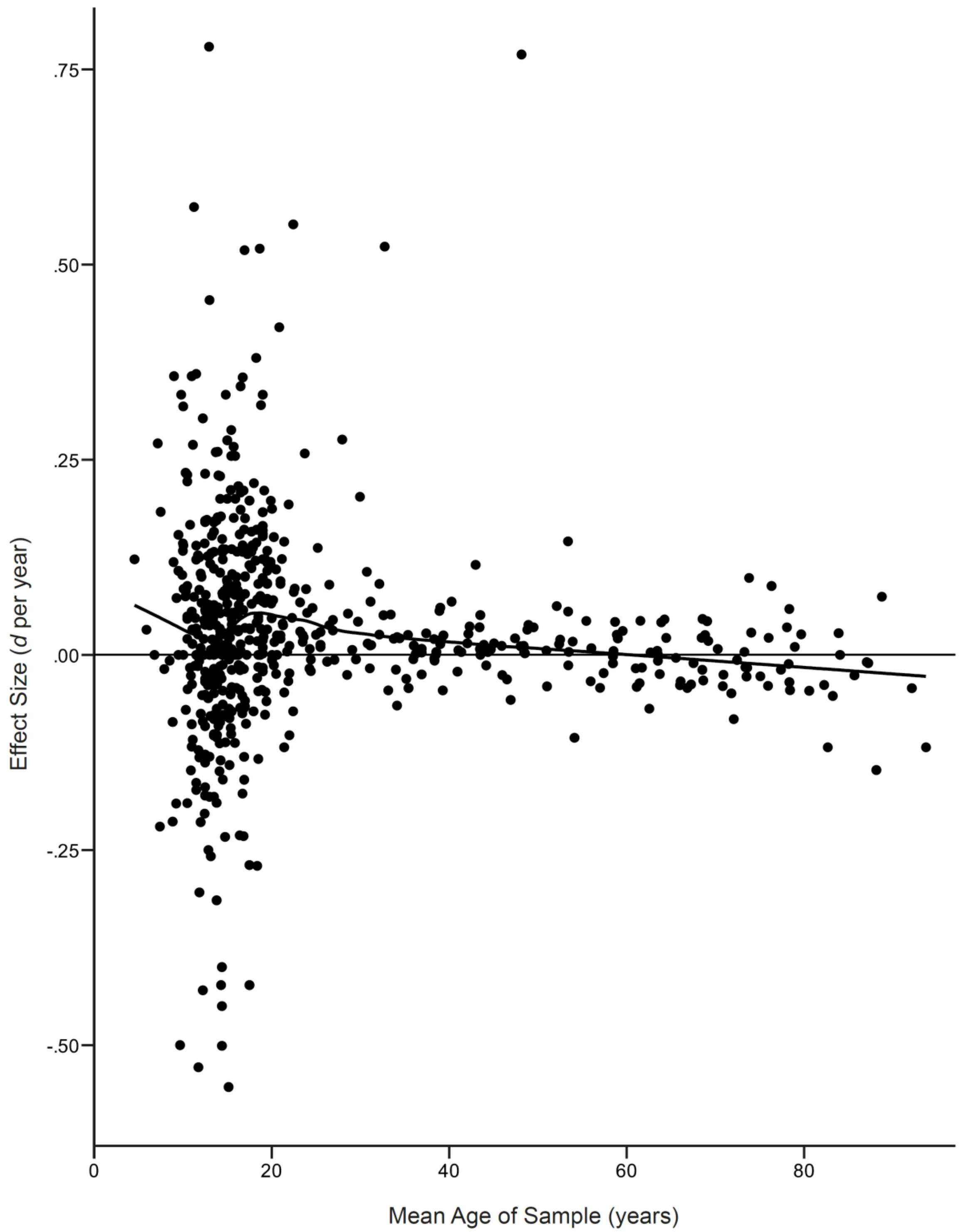

Figure 2. Scatterplot displaying the relation between effect size (i.e., standardized mean change *d* per year, d_{year}) and age (i.e., mean age of sample at the center of the observed time interval). The figure also shows the locally weighted smoothing (LOESS) curve across age.

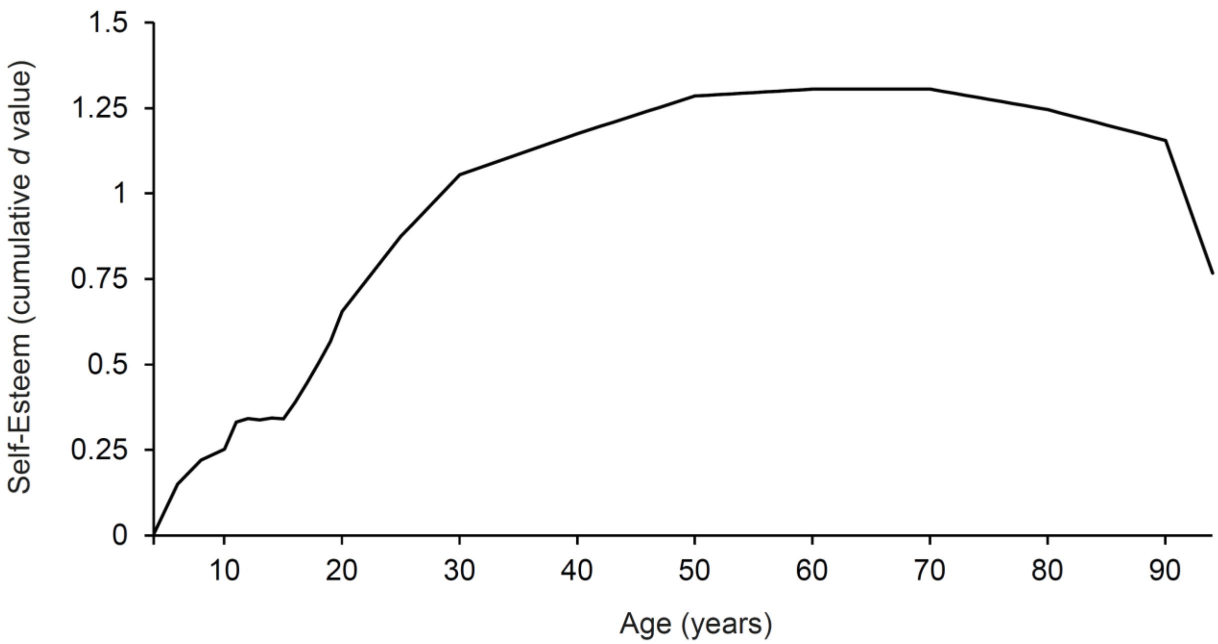

Figure 3. Mean-level change of self-esteem from age 4 to 94 years. The figure shows cumulative *d* values relative to age 4 years.