Journal of Research in Personality 70 (2017) 122–133

Contents lists available at ScienceDirect

Journal of Research in Personality

journal homepage: www.elsevier.com/locate/jrp

Full Length Article

High, low, and in between: Self-esteem development from middle childhood to young adulthood

Joanne M. Chung [a,*], Roos Hutteman [b], Marcel A.G. van Aken [b], Jaap J.A. Denissen [a]

[a] Department of Developmental Psychology, Tilburg University, Tilburg, Netherlands
[b] Department of Developmental Psychology, Utrecht University, Utrecht, Netherlands

ARTICLE INFO

Article history:
Received 1 March 2017
Revised 24 June 2017
Accepted 3 July 2017
Available online 4 July 2017

Keywords:
Self-esteem
Longitudinal
Shyness
Aggressiveness
Item response theory

ABSTRACT

We describe self-esteem development in a German sample (N = 240, 48% female) followed longitudinally from middle childhood to young adulthood, using data spanning 20 years. Data from the Self-Perception Profile for Children (Harter, 1985) and the Self-Description Questionnaire III (Marsh & O'Neill, 1985) were linked using item response theory methods. Rank-order stability was high in middle childhood, low in adolescence, and highest in young adulthood. Mean-levels were relatively high in middle childhood, decreased into adolescence, but increased into young adulthood. Early childhood shyness and aggressiveness as rated by parents, teachers, and observers did not influence the self-esteem trajectory. We provide the first longitudinal evidence for the self-esteem trajectory from middle childhood to young adulthood, replicating and extending previous findings.

1. Introduction

Global self-esteem refers to a person's subjective evaluation of his or her self-worth. The extent to which a person holds positive self-views has been shown to be important for fostering goals, coping strategies, and behaviors that enable success for one's relationships, career, and well-being (Kuster, Orth, & Meier, 2013; Orth, Robins, & Widaman, 2012; Steiger, Allemand, Robins, & Fend, 2014; Trzesniewski et al., 2006). Recently, there has been an influx of longitudinal research characterizing the trajectory of self-esteem (Orth, Maes, & Schmitt, 2015; Orth & Robins, 2014; von Soest, Wichstrøm, & Kvalem, 2016), offering much needed insights into the normative development of self-esteem from adolescence to young adulthood, and young adulthood to old age. Yet, to date, research on the lifespan trajectory of self-esteem has not included the shift from childhood to adolescence. During this time, children mature in their cognitive abilities, struggle with puberty, and become more concerned with their peers (Robins & Trzesniewski, 2005), making the transition from childhood to adolescence especially important for self-esteem development. Furthermore, there is little empirical research that has examined the developmental antecedents of self-esteem (but see Harris et al., 2015; Orth, 2017). Therefore, it is important to gain a better understanding of the lifespan development of self-esteem from middle childhood onward, and moreover, attempt to identify childhood factors that predict the trajectory of self-esteem itself.

In the present study, we examined global self-esteem development from middle childhood to young adulthood. We began by charting the trajectory of self-esteem, and sought to extend previous cross-sectional and longitudinal findings regarding self-esteem development by examining both stability and change during middle childhood, adolescence, and young adulthood. We used multiple informant reports (i.e., parent, teacher, and observer ratings) of early childhood personality to predict who increased or decreased in self-esteem during the transitions from childhood to adolescence and from adolescence to young adulthood. Specifically, we focused on shyness and aggressiveness between the ages of 4 and 6 years old as predictors of self-esteem because these traits have been linked to low self-esteem and difficulties in the social domain, raising the possibility that they could affect the development of self-esteem. Below, we review previous research on these topics, first turning to the literature on self-esteem development.

1.1. The development of self-esteem from middle childhood to young adulthood

Self-esteem is an evaluation of one's own worthiness and competence, but theoretical perspectives emphasize the importance of

* Corresponding author at: Department of Developmental Psychology, Tilburg University, PO Box 90153, 5000 LE Tilburg, Netherlands.
E-mail address: j.m.h.chung@uvt.nl (J.M. Chung).

http://dx.doi.org/10.1016/j.jrp.2017.07.001
0092-6566/

the social world in shaping self-esteem. Our self-views are thought to develop from our interactions with others and how we believe others see us (Cooley, 1902; Harter, 2012; Mead, 1934). Sociometer theory (Leary, Tambor, Terdal, & Downs, 1995) highlights the social nature of self-esteem and posits that it is a sociometer, or psychological gauge that signals the extent to which one is accepted by others, helping people maintain their social ties. The stable component of self-esteem is seen as one's judgment that he or she is generally valued and accepted by others, and as the "resting state" of the sociometer (Leary et al., 1995). Self-esteem is moderately stable across time and contexts, yet it is also mutable, especially during developmental transitions such as the ones from childhood to adolescence, and adolescence to young adulthood (Huang, 2010; Hutteman, Nestler, Wagner, Egloff, & Back, 2015). Both rank-order stability and mean-levels of self-esteem change across the lifespan. Rank-order stability refers to an individual's standing on the construct of interest, relative to others in the sample. Rank-order stability of self-esteem is lowest in early childhood and old age, relatively low but increasing in adolescence, and highest in adulthood (Donnellan, Kenny, Trzesniewski, Lucas, & Conger, 2012; Orth & Robins, 2014; Trzesniewski, Donnellan, & Robins, 2003). In childhood, stability is considered to be low because self-esteem is emerging and not fully formed during this time (Eccles, Wigfield, Harold, & Blumenfeld, 1993; Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002). In adolescence, stability is argued to be higher than in childhood due to an increased awareness of self, but lower relative to young adulthood because of maturational and social changes that are experienced during this time (Orth & Robins, 2014; Robins & Trzesniewski, 2005). Accordingly, in the present study, we expected to find low rank-order stability in childhood, low but increasing stability into adolescence, and high stability into young adulthood.

Mean, or average levels of self-esteem, are relatively high in childhood, decrease during adolescence, and then steadily increase into young adulthood (Orth, Trzesniewski, & Robins, 2010; Orth et al., 2012, 2015; Robins et al., 2002; von Soest et al., 2016). Young children are thought to exhibit relatively high self-esteem because their self-views are unrealistically positive (Robins & Trzesniewski, 2005). As children's cognitive skills mature, they begin to compare their skills and abilities to their peers and consider feedback from close others, including the extent to which they are liked and accepted by others, and develop more accurate, and generally, less positive self-views (Robins & Trzesniewski, 2005). Self-esteem continues to decrease into adolescence, with pubertal changes, changing school contexts, and the increased capacity for self-reflection being implicated in this decline (Orth & Robins, 2014; Robins & Trzesniewski, 2005). During the transition into young adulthood, self-esteem has been found to increase as individuals increase in autonomy at school and work, and deepen their social relationships with others (Orth & Robins, 2014). Accordingly, in the present study, we expected to find self-esteem to be relatively high in mid-childhood, decrease into adolescence, and then increase into young adulthood.

Self-esteem development has been examined almost exclusively in childhood (Arunkumar, Midgley, & Urdan, 1999; Rodriguez, Wigfield, & Eccles, 2003), from childhood to adolescence (Hoge, Smit, & Hanson, 1990; Wigfield, Eccles, Mac Iver, Reuman, & Midgley, 1991), or from adolescence to adulthood (Erol & Orth, 2011; Orth & Robins, 2014; von Soest et al., 2016), including in a recent study that examined self-esteem in the LOGIC data from when youth were 17- to 29-years-old (Luan et al., 2017). These studies have contributed much towards our understanding of self-esteem development. Yet, when piecing together the evidence from different developmental and time periods, we cannot rule out cohort effects – the possibility that age differences in self-esteem are confounded with experiences of events that are not shared with other populations (Baltes, Cornelius, & Nesselroade, 1979). For example, some have claimed that societal shifts in the focus on the self have influenced the nature and developmental course of self-esteem (Twenge, 2006). Studies linking childhood to adulthood are rare, understandably, because researchers are often faced with the issue of heterotypical continuity, where manifestations of the same underlying trait change as individuals mature. In the present study, we had the unique opportunity to track self-esteem in the same group of individuals over two decades, allowing us to examine stability and change in self-esteem from middle childhood to young adulthood.

1.2. Early childhood shyness and aggressiveness as antecedents of self-esteem development

Personality is thought to influence the ways individuals construe the world around them (Caspi & Shiner, 2011; Rothbart, 2011) as well as themselves (Robins, Donnellan, Widaman, & Conger, 2010; Robins, Tracy, Trzesniewski, Potter, & Gosling, 2001). Specifically, personality can influence which features of a social interaction that one attends to, in turn, coloring one's perceptions and experience, including one's self-evaluations (Caspi & Shiner, 2011). Shyness and aggressiveness are aspects of personality that reveal how individuals relate to the social world, and reflect both temperament and social competence (Caspi, Elder, & Bem, 1987, 1988; Denissen, Asendorpf, & Van Aken, 2008; Horney, 1950; Rubin, Coplan, & Bowker, 2009). Shyness refers to a tendency towards experiencing worry, tension, or awkwardness during social interactions, especially with strangers (Cheek & Buss, 1981; Coplan & Rubin, 2010). Aggressiveness, in contrast, refers to a proneness towards behaving with the intention of causing harm to another person (Anderson & Bushman, 2002). Caspi et al. (1987, 1988) operationalized "moving away from the world" and "moving against the world" as informant ratings of children's tendencies towards exhibiting shy and aggressive behaviors respectively, and found that being perceived as either shy or aggressive in childhood predicted poor outcomes in the psychological, social, and work domains 30 years later. Shyness and aggressiveness in childhood are thought to have such powerful consequences on individuals' life trajectories because of their consistency throughout the lifespan (Caspi et al., 1987, 1988).

The extant research indicates that shyness in childhood is concurrently associated with low self-esteem (Crozier, 1995; Kemple, David, & Wang, 1996; Rubin et al., 2009). Studies also indicate that shyness in childhood shows negative prospective effects for self-esteem. In a previous study that examined the LOGIC data, youth who were seen as extremely shy towards familiar peers at 4- to 6-years-old were likely to report low levels of self-esteem between the ages of 8- and 10-years-old (Asendorpf & van Aken, 1994). In another, longitudinal study, Icelandic youth who were categorized as shy at 7-years-old were likely to exhibit low levels of self-esteem in early and middle adolescence (Hart, Hoffmann, Edelstein, & Keller, 1997). Shyness in childhood is also predictive of a restricted social life (Gest, 1997) and low self-esteem (Rubin, Chen, McDougall, Bowker, & McKinnon, 1995) in adulthood. Additionally, Swedish girls who were perceived as shy in early childhood were likely to exhibit low levels of self-esteem 30 years later (Kerr, 2000). These results suggest that we might expect that individuals perceived as shy in early childhood might be negatively affected in their self-esteem development.

With regard to aggressiveness, the extant research suggests that it is associated with low self-worth (Coplan, Prakash, O'Neil, & Armer, 2004; Donnellan, Trzesniewski, Robins, Moffitt, & Caspi,

2005; but see Denissen, Thomaes, & Bushman, in press). And, like shyness, longitudinal studies that examine aggressiveness in childhood suggests that it has negative prospective effects on constructs related to self-esteem. In one longitudinal study, American children who were rated by teachers as aggressive in kindergarten were less likely to be accepted by their peers, more likely to have fewer friends, and more likely to be lonelier than their less aggressive peers from kindergarten to the 2nd grade (Ladd & Burgess, 1999). In another, longitudinal study, Norwegian adolescents who were seen by their mothers as aggressive from infancy to mid-adolescence were likely to report low levels of life satisfaction and high levels of depression and anxiety (Kjeldsen et al., 2016). In yet another, longitudinal study, American children who were seen as aggressive at 7-years-old were likely to experience sharply increasing levels of loneliness from late childhood to mid-adolescence (Schinka, van Dulmen, Mata, Bossarte, & Swahn, 2013).

Taken together, while much of the research on self-esteem has focused on its prospective effects (e.g., Kuster et al., 2013; Orth et al., 2012; Trzesniewski et al., 2006), including on aggressiveness (Denissen et al., in press; Donnellan et al., 2005), conceptual and empirical links point to the possibility that children who possess traits that might hinder successful social relations, such as high levels of shyness and aggressiveness, might be negatively impacted in their subsequent self-esteem development. Examining this possibility not only provides a longitudinal consideration of the ideas posited by sociometer theory (Leary et al., 1995), but can shed light on whether or not early childhood is an appropriate time to target these ways of interacting with the world. To date, however, there is no research examining how shyness and aggressiveness in early childhood predicts changes in self-esteem from middle childhood to adulthood.

1.3. Present study

In the present study, we examined global self-esteem development in a sample of individuals tracked over 20 years. Self-esteem was assessed using two different measures of global self-esteem. In order to chart the trajectory of self-esteem from middle childhood to young adulthood, we linked scores from the two different measures of global self-esteem using item response theory methods. Following previous cross-sectional and longitudinal findings regarding self-esteem development (Orth & Robins, 2014; Robins et al., 2002), we expected that self-esteem would be relatively high in childhood, decrease into adolescence, and then increase into young adulthood. We then tested whether or not shyness and aggressiveness in early childhood influenced the trajectory of self-esteem. We expected that those who were seen as having shy or aggressive personalities in childhood would show a less positive self-esteem trajectory relative to their peers.

In the present study, we adopted a multi-method approach in addressing our research aims. Specifically, we used data from multiple informants, consisting of parent, teacher, and observer ratings of early childhood shyness and aggressiveness. These data consisted of measures that differed in their objectivity – ranging from the familiar perceptions of the participant's shyness and aggressiveness made by parents and teachers to the more distal judgments of observers in both laboratory and field settings. By using this approach to assessing early childhood personality, we were able to obtain a more comprehensive perspective of shyness and aggressiveness and their influences on participant ratings of their own self-esteem, moving beyond previous research that has relied on either self- or single informant ratings.

2. Method

2.1. Sample and procedure

We used data from the Munich Longitudinal Study on the Genesis of Individual Competencies (LOGIC),[1] a longitudinal study that was, in part, designed to examine the effects of personality and social experiences on adjustment (for further details about the study, see Schneider & Bullock, 2009; Weinert & Schneider, 1999). Our sample consists of 240 individuals (48% female) who were born in 1980 or 1981 and who attended preschool in Munich, Germany. Families were recruited from 20 preschools, and more than 90% allowed their child to participate in the study. All children spoke German as their first language, and according to parent reports of professional status (Wegener, 1988), 63% of families were of middle SES, 28% were of low SES, and 9% were of high SES.

The present study used 9 waves of data, when youth were 4, 5, 6, 9, 10, 12, 17, 23, and 29 years of age. Assessments for the first 7 assessments were conducted at the Max Planck Institute for Psychological Research in Munich, and assessments for the last two waves were administered in an online questionnaire. Families were contacted several times during the year to encourage participation in the study; retention rates were relatively high from the Age 4 to Age 12 assessments, and decreased slightly from the Age 17 to Age 29 assessments (see Supplemental Information, p. 1, Table 1 for assessment intervals and sample sizes for each type of rating). To investigate the impact of attrition, we compared individuals who provided self-esteem ratings at both the Age 9 and Age 29 assessments to those who did not provide self-esteem ratings at the Age 29 assessment on the study variables. Those who provided self-esteem ratings at both waves were rated as lower on the aggregated Shyness Towards Unfamiliar Peers subscale as rated by parents ($M = 2.86$) (described in the Measures section below) than those who did not provide self-esteem ratings at the Age 29 assessment ($M = 3.65$), $F(1, 95) = 8.51$, $p < .05$, $\eta_p^2 = .08$). We did not find any additional effects of attrition on any of the other main study variables.

2.2. Measures

2.2.1. Self-esteem

Two measures of self-esteem were administered at six assessments (i.e., Age 9, 10, 12, 17, 23, and 29), and did not overlap. In the Age 9 to Age 12 assessments, a 6-item German version of the Global Self-Worth subscale from the Self-Perception Profile for Children (SPPC; Harter, 1985) was administered. Children were first asked which part of the statement was most like them (e.g., "Some children are pleased with themselves but other children are not pleased with themselves"), and then asked how well it described them. Responses were then coded from 1 (very low self-esteem) to 4 (very high self-esteem). Coefficient alphas ranged between .76 and .77. In the Age 17 to Age 29 assessments, a 6-item German version of the General Self-Concept subscale from the Self-Description Questionnaire III (SDQ-III; Marsh & O'Neill, 1985) was

[1] Data from the LOGIC study have been used in several studies. Studies focused on social development have examined shyness (Asendorpf, 1992), as well as aggressiveness, including their developmental trajectories (Denissen et al., 2008), long-term stability and effect on life outcomes (Asendorpf, Denissen, & van Aken, 2008), the interplay between the two (Hutteman, Denissen, Asendorpf, & van Aken, 2009), and their associations with the Big Five personality traits in childhood (Asendorpf & van Aken, 2003). Additionally, some studies have focused on self-esteem during late adolescence and young adulthood (Luan et al., 2017; Sturaro, Denissen, van Aken, & Asendorpf, 2008), as well as the extent to which shyness predicts social self-esteem during childhood (Asendorpf & van Aken, 1994). However, no studies have examined the trajectory of global self-esteem development from childhood to young adulthood, and none have examined how early childhood shyness and aggressiveness predict individual differences in this trajectory.

administered (e.g., "All in all, I accept myself as I am"). Response options for this measure ranged from 1 (*not at all*) to 5 (*completely*). Coefficient alphas ranged between .79 and .80.

2.2.2. Self-esteem-related items

In the Age 12 and Age 17 assessments, bipolar adjective pairs assessing Big Five personality dimensions (Ostendorf, 1990) were administered. We identified two items measuring the Emotional Stability dimension ("self-confident – vulnerable", and "insecure" – "self-confident") that are related to self-esteem (*r*s among the items and the self-esteem score were both .21 at Age 12, and .40–.50 at Age 17; all *p*s < .05). Participants evaluated each adjective pair in relation to each other, such that response options ranged from 1 (*very [first adjective], not [second adjective]*), to 5 (*not [first adjective], very [second adjective]*). The responses to the first item was reverse-coded.

2.2.3. Shyness and aggressiveness in early childhood

Six measures of shyness and three measures of aggressiveness as rated by three informants were examined for the Age 4 to Age 6 assessments. As in previous research on the LOGIC study (Asendorpf & van Aken, 1999), averages were computed across these three assessments to represent indices of shyness and aggressiveness in early childhood, with the exception of observer ratings derived from the laboratory task, for which Age 5 data were not available.

2.2.3.1. Parent ratings. Parents completed a questionnaire designed to assess their perceptions of their child's shyness and aggressiveness. The 4-item Shyness Towards Unfamiliar Adults (SUA) subscale (e.g., "My child is shy towards unknown adults") and the 4-item Shyness Towards Unfamiliar Peers (SUP) subscale were examined. The items in the subscales were identical, but referenced either other children or adults. Parents also completed the 4-item Aggressiveness Towards Peers (AP) subscale (e.g., "My child is aggressive towards other children"). Coefficient alphas showed ranges between .93 and .95, .89 and .94, and .82 and .85 for SUA, SUP, and AP, respectively.

2.2.3.2. Teacher ratings. Teachers completed a 54-item German version of the California Child Q-Set (Göttert & Asendorpf, 1989) to assess the children's personality. Teachers were instructed to assign six items each into nine categories, ranging from 1 (*extremely uncharacteristic*) to 9 (*extremely characteristic*). Prototypicality scores were derived by correlating each child's Q-sort profile with a prototypic Q-sort profile for shyness (e.g., "The child is inhibited and constricted") and aggressiveness (e.g., "The child is aggressive"). Coefficient alphas were above .80 and .78 for the three assessments of shyness and aggressiveness, respectively.

2.2.3.3. Observer ratings. Children were observed interacting in a laboratory task meant to elicit behavioral inhibition, and free play at preschool (see Asendorpf, 1990 for more information). For shyness, Contact Initiation Latency (the number of seconds it took for the child to approach the stranger) and Shyness Towards Strangers ratings (how shy the child seemed to be in presence of the stranger) were obtained from the laboratory task. For the latter, observer ratings ranged from 1 (*not at all*) to 7 (*extremely*), and coefficient alphas ranged between .89 and .97. Wait-and-Hover ratings (the percentage of social interactions initiated by child after gaining proximity to other children relative to the total number of initiations) were obtained from the free play task, and kappa reliability was above .90. For aggressiveness, Aggressive Attacks among All Own Initiations ratings (the percentage of verbal and physical acts of aggression relative to the total number of initiations) were obtained from the free play task, and kappa reliability was above .90.

2.3. Power considerations for the current study

The literature regarding adequate power for item response theory models does not provide a definitive answer. Some researchers recommend 250 or 500 respondents for the Graded Response Model (Embretson & Reise, 2000), while others say that this recommendation may be relaxed if there are reasonable associations between the items with the latent factor (K.J. Grimm, personal communication, May 16, 2017) (see Supplemental Information, p. 2, Table 2 for factor loadings). For structural equation models, some researchers recommend a minimum sample size of 100 or 200 (Boomsma, 1982, 1985; Kline, 2005), while others recommend 5 or 10 observations per estimated parameter (Bentler & Chou, 1987). Thus, readers should interpret our findings under the caveat that our sample (N = 195 for the unconditional model) is modestly powered depending on which rule of thumb the reader favors. With regard to the structural equation models, we conducted Montecarlo simulations to examine post-hoc power for each of the parameters estimated by the growth models we conducted, and to take into account missing data. We did this by entering the estimates provided by our models as population values, including the patterns of missingness we observed in our data and the proportion of subjects within each pattern. We requested 10,000 replications for each of our models.

2.4. Description of linking procedure using item response theory methods

The goals of using a linking procedure were to describe the trajectory of self-esteem from when the child was 9-years-old to when they were 29-years-old (see Fig. 1 for original scale scores).

We employed item response theory methods to link two different measures of self-esteem to each other (Curran et al., 2008; de Ayala, 2009; Kolen & Brennan, 2004). Linking is a less restrictive form of equating, a group of procedures that refer to the adjustment of person location estimates derived from different forms to place them on a common metric (de Ayala, 2009; Kolen & Brennan, 2004).

Because there was no overlap in the administration of self-esteem measures in the Age 12 and Age 17 assessments, we used external common items – items that are closely related to the constructs of interest and that were administered during both Age 12 and Age 17 assessments. External common items do not contribute directly to the final estimates, but are used to calibrate, or place the two different measures of a construct on the same metric (de Ayala, 2009; Kolen & Brennan, 2004). In the present study, we linked the global self-esteem data from the SPPC (Harter, 1985) and SDQ-III (Marsh & O'Neill, 1985) using the self-esteem related items described above.

We took a concurrent calibration approach (de Ayala, 2009; Kolen & Brennan, 2004) to linking the data. This method treats item characteristics (i.e., difficulty and discrimination) for the common items on one form (e.g., the data from the Age 12 assessment) as fixed, true values, and then scales the items on the other form (e.g., the data from the Age 17 assessment) according to the estimates of the fixed common items (de Ayala, 2009; Kolen & Brennan, 2004). Under this approach, scaling was done simultaneously across all times of measurement (de Ayala, 2009; Kolen & Brennan, 2004).

For each construct, we began by fitting a Graded Response Model (GRM; Samejima, 1969) that included the items from the first measure of the construct at the Age 12 assessment, the two external common items at the Age 12 and Age 17 assessments,

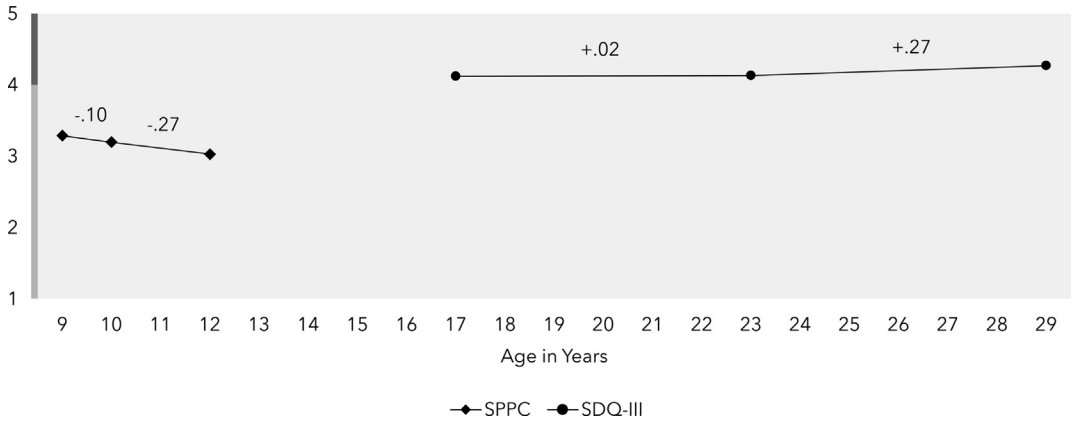

Fig. 1. Self-esteem from Age 9 to 29 as assessed by the self-perception profile for children and Self-Description Questionnaire-III. *Note.* $N = 195$. Effect sizes (Cohen's d) are shown in the figure above. Light grey axis denotes range of response categories for the Self-Perception Profile for Children (SPPC; Harter, 1985); dark grey axis denotes additional response category for the Self-Description Questionnaire III (SDQ-III; Marsh & O'Neill, 1985).

and the items from the second measure of the construct at the Age 17 assessment. The GRM is an item response model that estimates both discrimination (e.g., how well an item differentiates people of different standing on the latent trait, theta) and location parameters (e.g., how difficult an item is to endorse given one's standing on theta) for items that have more than two response categories (Samejima, 1969). We saved the estimated item parameters from this GRM.

We continued by fitting another GRM to all of the longitudinal data available from the Age 9 through Age 29 assessments in one step, using the saved parameters from the previous GRM as fixed estimates in which to calibrate items from the two different measures onto the same metric. By linking the data from the two different measures using this method, the model provided estimates of theta – the underlying latent traits – in our case, self-esteem – for each individual at each wave, allowing us to examine the trajectory of self-esteem from when individuals were 9-years-old to when they were 29-years-old. We then saved the theta estimates as individual factor scores, and used these values in subsequent analyses (see Fig. 2 for these estimates). As is generally the case in this procedure, theta was scaled so that it had a mean of 0 and a standard deviation of 1 (de Ayala, 2009; Kolen & Brennan, 2004).

3. Results

We conducted our study analyses using the psych Version 1.5.8 (Revelle, 2015) and mirt (Chalmers, 2012) packages in R Version 3.2.3 (R Core Team, 2015), and Mplus Version 7 (Muthén & Muthén, 1998–2012). We used the Root Mean Square Error of Approximation (RMSEA) ($\leq$ .08), the Comparative Fit Index (CFI) ($\geq$ .90), and the Tucker-Lewis Index (TLI) ($\geq$ .90) as indices to determine adequate model fit (Browne & Cudeck, 1993; Hu & Bentler, 1999; see Table 1 for fit statistics for all models and Table 2 for descriptive statistics).

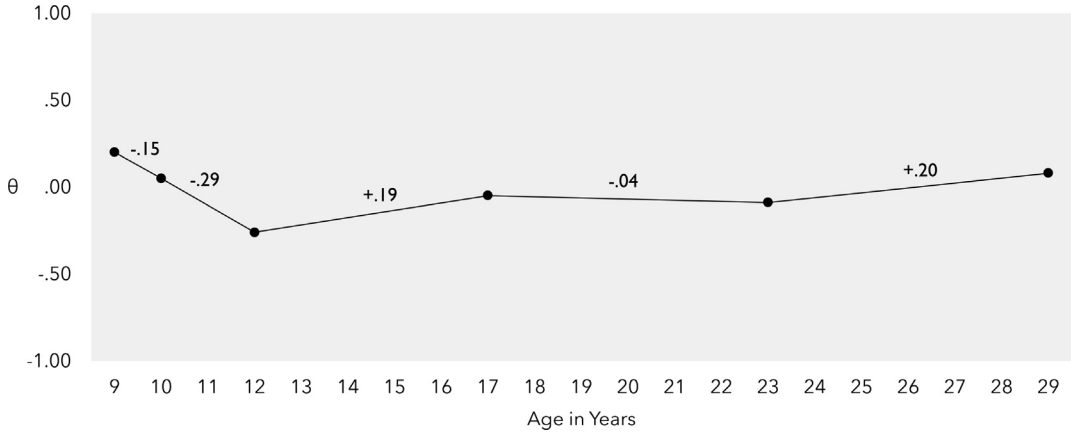

Fig. 2. Self-esteem from Age 9 to 29 as estimated by linking using item response theory methods. *Note.* $N = 195$. Theta represents latent self-esteem and is scaled so that it has a mean of 0 and a standard deviation of 1. Effect sizes (Cohen's d) are shown in the figure above.

Table 1
Fit statistics for all models.

No.	Description	*RMSEA*									
		N	χ^2	*df*	*p*	est.	90% C.I.		*p*	*CFI*	*TLI*
1	First-order autoregressive for self-esteem	195	6.325	9	.707	.000	.000	.061	.907	1.000	1.000
2	Unconditional for self-esteem	195	12.307	11	.341	.025	.000	.081	.705	.990	.986
3	Shyness and self-esteem	232	28.149	28	.457	.005	.000	.051	.944	.999	.999
4	Aggressiveness and self-esteem	233	37.478	28	.109	.038	.000	.067	.717	.945	.929

Table 2
Descriptive statistics for predictor variables.

	1.	2.	3.	4.	5.	6.	7.	8.	9.
1. Shyness towards adults (parent)	1.000								
2. Shyness towards peers (parent)	.802*	1.000							
3. Shyness Q-sort (teacher)	.361*	.459*	1.000						
4. Shyness towards strangers (observer)	.651*	.521*	.466*	1.000					
5. Wait-and-hover (observer)	.179	.222	.484*	.276*	1.000				
6. Contact initiation latency (observer)	.530*	.511*	.273*	.443*	.080	1.000			
7. Aggressiveness towards peers (parent)	−.165	−.062	−.274*	−.152	−.147	−.095	1.000		
8. Aggressiveness Q-sort (teacher)	−.204*	−.267*	−.661*	−.406*	−.264*	−.097	.271*	1.000	
9. Aggressive attacks among all own initiations (observer)	−.160	−.287*	−.421*	−.128	−.383*	.027	.215*	.441*	1.000
Mean	3.828	3.104	.035	3.435	21.619	330.224	2.818	−.213	2.765
SD	1.296	1.116	.277	1.243	12.153	249.945	.785	.348	4.364

Note. Correlation table computed using all available data in Mplus ($N = 231$). All variables are aggregated across Age 4, Age 5, Age 6 assessments, with the exception of observer ratings of Shyness Towards Strangers and Contact Initiation Latency, which did not have an Age 5 assessment. Wait-and-Hover and Contact Initiation Latency are in seconds.
* $p < .05$.

Table 3
Estimates for first-order autoregressive model (Model 1).

Regressions	Age range	Interval	Parameter estimates						
			Original				Annualized		
			B	*SE B*	*b*	*p*	Power**	*B*	*b*
S2 θ on S1 θ	9–10	1	.820	.196	.841	<.001	.992	.820	.841
S3 θ on S2 θ	10–12	2	.494	.144	.570	.001	.981	.703	.755
S4 θ on S3 θ*	12–17	5	.531	.184	.427	.004	.837	.881	.844
S5 θ on S4 θ	17–23	6	.779	.178	.742	<.001	.998	.959	.951
S6 θ on S5 θ	23–29	6	.834	.157	.965	<.001	1.000	.970	.994

Note. $N = 195$. S denotes the latent status indicated by the manifest theta estimate. Age range and interval are in years. Parameter estimates are annualized for time interval by taking the *n*th root of the estimate, where *n* is equal to the length of the interval (i.e., number of years between assessments).
* The self-esteem measure changed between S3 and S4. Parameter estimates for full model are shown in Supplemental Information, Table 4.
** A Montecarlo simulation was conducted to determine the power of each estimated parameter model output entered as population values; it is possible to obtain a non-positive definite model implied variance-covariance matrix, which means that the standard errors of the parameter estimates are not estimable (a common example of this would be if predictors in a regression are too highly correlated), and this occurred .15% of the time. For significant effects, we expected power to be high (e.g., the regression of S2 on S1); for non-significant effects, we expected power to be low (e.g., the covariance between the childhood and youth slopes).

3.1. Rank-order stability of self-esteem from middle childhood to young adulthood

To examine the rank-order stability of self-esteem from childhood to young adulthood, we fit a first-order autoregressive model (Model 1) to the theta estimates for the six assessments, fixing the measurement error to be equal across assessments (see Table 3 for model estimates and a description of how estimates were corrected for time interval). First-order autoregressive models are used on repeated measures data and are in contrast to traditional Ordinary Least Squares methods because they take into account the dependency of an earlier assessment on later assessments, such that each assessment consists of the previous assessment and a random error component. Therefore, first-order autoregressive models provide estimates of the latent variable separately from error, allowing for stability estimates that have been corrected for unreliability of the measure.

Rank-order stability for self-esteem was lowest between the Age 10 and Age 12 assessments (*b* corrected for time interval = .755), and highest between the Age 23 and Age 29 assessments (*b* corrected for time interval = .994).

3.2. Mean-level change in self-esteem from middle childhood to young adulthood

Our next step was to examine mean-level change and individual differences in the trajectory of self-esteem. A visual inspection of the factor scores associated with the theta estimates provided by our linking procedure (see Fig. 2 for theta estimates) suggested a non-linear trajectory, with a decrease at the Age 12 assessment. Thus, we fit a piecewise growth model. We first fit an unconditional model (Model 2) to determine whether there was an average, or typical, trajectory of self-esteem. In addition to a latent intercept, we specified two latent slopes that were centered at the Age 9 assessment, with these coefficients being 0. For the first latent slope, the five remaining assessments were given values of 1, 3, 3, 3, 3, corresponding to the difference in years between assessments for the first piece (i.e., the number of years between the Age 9 and Age 10 assessments was 1, the number of years between the Age 9 and Age 12 assessments was 3, and to accommodate the specification of the second piece, the remaining assessments were fixed at 3). For the second latent slope, the Age 10 and Age 12 assessments were both given a value of 0, reflecting the fact that the specification of the second piece began at the Age 17 assessment. The Age 17 assessment was allowed to be freely estimated, and the Age 23 and Age 29 assessments were given values of 11 and 17, corresponding to the difference in years between assessments after accounting for the fixed value of 3 for the same assessments in the first piece (i.e., the number of years between the Age 9 and Age 23 assessments, and between the Age 9 and Age 29 assessments were 14 and 20 years, respectively).[2]

[2] To make sure that there were no other, simpler models that could adequately characterize our data, we fit additional models: linear, latent basis, quadratic, cubic, and piecewise linear models. Each of these models differ in the way the latent slope(s) were specified. See Supplemental Information, pp. 3–4, Table 3 for model specifications and fit. The results of these models indicate that the piecewise latent basis model was the most reasonable model for our data.

By specifying the latent slopes this way, we created two "pieces" that converged at the self-esteem drop that occurred at the Age 12 assessment. Additionally, these values take into account the uneven spacing of assessments (e.g., the length of time between assessments become much greater in the later assessments than in the beginning) and the slight non-linearity at the Age 17 assessment. Thus, we found that a model with one latent intercept, one latent linear slope, and one latent basis slope that were allowed to covary fit the data well (see Fig. 3 for path diagram).

Below we refer to the latent intercept as the *self-esteem intercept*, the first latent slope as the *childhood slope* (covering the Age 9 to Age 12 assessments, and the second latent slope as the *youth slope* (covering the Age 12 to Age 29 assessments) (see Table 4 for model estimates).

The mean (B = .199) and variance (B = .408) of the self-esteem intercept were significant, indicating that on average, children started out with relatively high levels of self-esteem, but that there were individual differences in initial levels of self-esteem. The mean of the childhood slope was negative and also significant (B = −.153), but its variance was not significant (B = .031), indicating that on average, children decreased in self-esteem into late childhood, and also, that the majority of the youth in our sample experienced this decrease. Due to the non-significant variance of the childhood slope, we do not interpret the associations between this latent variable and others in this model and in subsequent models that include it. The mean (B = .019) and variance (B = .002) of the young adulthood slope were significant, indicating that on average, youth increased in self-esteem from adolescence to young adulthood and also, that there were individual differences in this trajectory. The self-esteem intercept was not significantly associated with the young adulthood slope (r = −.286), indicating that the increase in self-esteem into young adulthood was not associated with self-esteem levels when children were 9-years-old.

Table 4
Estimates for model of self-esteem trajectory (Model 2).

	B	*SE B*	*b*	*p*	Power*
Model					
Youth slope (Y S) by					
θ Age 9	0	–	–	–	–
θ Age 10	0	–	–	–	–
θ Age 12	0	–	–	–	–
θ Age 17	9.279	2.214	–	<.001	.991
θ Age 23	11	–	–	–	–
θ Age 29	17	–	–	–	–
Covariances					
Childhood slope (CS) with self-esteem intercept (S-EI)	−.078	.032	−.693	.014	.693
Y S with S-EI	−.008	.004	−.286	.070	.441
Y S with C S	.000	.003	.059	.864	.059
Means					
S-E I	.199	.060	–	.001	.913
C S	−.153	.025	–	<.001	1.000
Y S	.019	.005	–	<.001	.963
Variances					
S-E I	.408	.088	–	<.001	.999
C S	.031	.017	–	.064	.422
Y S	.002	.001	–	.008	.847

Note. N = 195. θ represents latent trait of self-esteem as estimated by linking procedure. Parameter estimates for full model are shown in Supplemental Information, Table 5.

* A Montecarlo simulation was conducted to determine the power of each estimated parameter model output entered as population values; it is possible to obtain a non-positive definite model implied variance-covariance matrix, which means that the standard errors of the parameter estimates are not estimable (a common example of this would be if predictors in a regression are too highly correlated), and this occurred .01% of the time. For significant effects, we expected power to be high (e.g., the mean of the self-esteem intercept); for non-significant effects, we expected power to be low (e.g., the covariance between the childhood and youth slopes). Dashed lines indicate that parameters were not estimated or that standardized estimates are not presented.

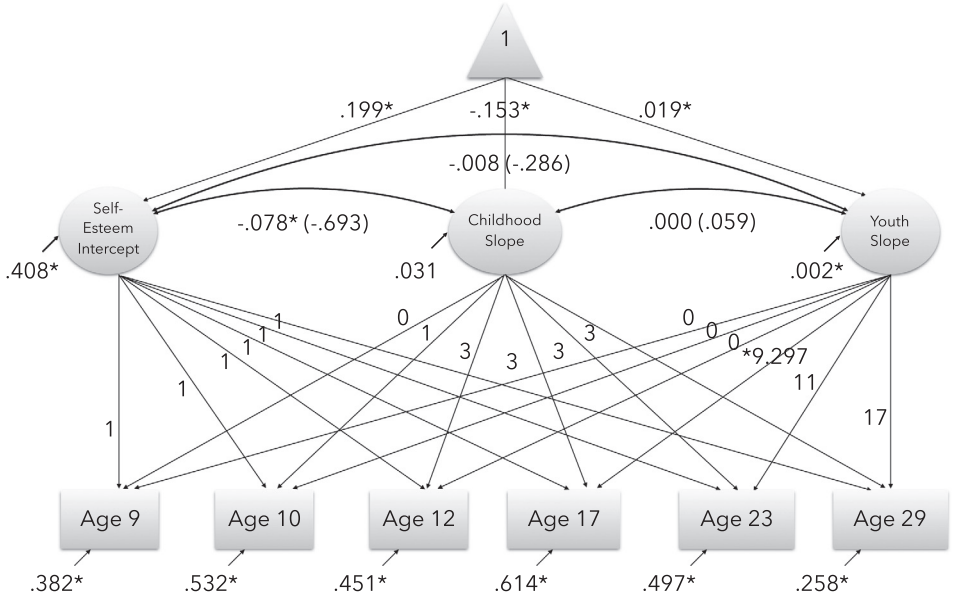

Fig. 3. Path diagram for model of self-esteem trajectory (Model 2). *Note.* N = 195. We specified two latent slopes centered at the Age 9 self-esteem assessment, and assigned a value of 0 to these coefficients. For the first latent slope, the five remaining assessments were given values of 1, 3, 3, 3, 3, corresponding to the difference in years between assessments for the first piece (i.e., the number of years between the Age 9 and 10 assessments was 1, the number of years between Age 9 and Age 12 was 3, and to accommodate the specification of the second piece, the remaining assessments were fixed at 3). For the second latent slope, the Age 10 and Age 12 assessments were both given a value of 0, reflecting the fact that the second piece began at the Age 17 assessment. The Age 17 assessment was allowed to be freely estimated, and the Age 23 and Age 29 assessments were given values of 11 and 17, respectively, corresponding to the difference in years between assessments after accounting for the fixed value of 3 for the same assessments in the first piece (i.e., the number of years between the Age 9 and Age 23 assessments, and between the Age 9 and Age 29 assessments were 14 and 20 years, respectively). Standardized estimates are presented in parentheses. * p < .05, two-tailed.

Table 5
Shyness and aggressiveness as predictors of self-esteem trajectory (Model 3 and Model 4).

Parameter	Shyness (Model 3)					Parameter	Aggressiveness (Model 4)				
	B	*SE B*	*b*	*p*	Power*		*B*	*SE B*	*b*	*p*	Power*
Regressions						Regressions					
Self-esteem intercept on shyness	−.062	.182	−.044	.732	.060	Self-esteem intercept on aggressiveness	.231	.193	.176	.231	.238
Childhood slope on shyness	−.046	.070	−.120	.513	.101	Childhood slope on aggressiveness	−.156	.097	−.420	.109	.548
Youth slope on shyness	.001	.015	.014	.931	.048	Youth slope on aggressiveness	.012	.015	.137	.428	.126

Note. *Ns* = 232 and 233 for Model 3 and Model 4, respectively. Models were estimated separately. Parameter estimates for full model are shown in Supplemental Information, Tables 6 and 7.

* Montecarlo simulations were conducted to determine the power of each estimated parameter model output entered as population values; it is possible to obtain a non-positive definite model implied variance-covariance matrix, which means that the standard errors of the parameter estimates are not estimable (a common example of this would be if predictors in a regression are too highly correlated), and this occurred .1% and 3.1% of the time for Model 3 and Model 4, respectively. For significant effects, we expected power to be high; for non-significant effects, we expected power to be low.

3.3. Early childhood shyness and aggressiveness as antecedents of the self-esteem trajectory

Do shyness and aggressiveness as rated by parents, teachers, and observers of youth in early childhood predict individual differences in the trajectory of self-esteem from middle childhood to young adulthood? To address this question, we specified separate conditional models (Model 3 and Model 4) to examine the potential moderating effects of shyness and aggression on the self-esteem trajectory (see Table 5 for model estimates).

Parent, teacher, and observer ratings for shyness and aggressiveness were first standardized and then entered into each model as manifest indicators of a latent variable. In the case of more than one variable for each informant, the average of the standardized variables was used. Specifically, to test the effects of shyness (Model 3), we used the average of parent ratings as assessed by the SUA and SUP subscales, teacher ratings of shyness, and the average of observer ratings as measured by Contact Initiation Latency, Shyness with Strangers, and Wait-and-Hover as manifest indicators of a latent Shyness factor. To test the effects of aggressiveness (Model 4), we used the parent ratings as assessed by the AP subscale, teacher ratings of aggressiveness, and observer ratings as assessed by Aggressive Attacks among All Own Initiations as manifest indicators of a latent Aggressiveness factor.

For both models, our results show that there were no significant effects on the self-esteem trajectory, indicating that neither shyness nor aggressiveness as rated by parents, teachers, and observers when individuals were 4- to 6-years-old predicted self-esteem development when individuals were 9-years-old to 29-years-old.

4. Discussion

In the present research, we examined global self-esteem development in a sample of individuals followed longitudinally from age 9 to age 29, and then tested the possibility that children's styles of relating to the social world in early childhood – as perceived by parents, teachers, and observers – were predictive of individual differences in the trajectory of self-esteem. Our findings contribute to the literature on self-esteem development by describing its longitudinal trajectory from middle childhood, which is in contrast to previous studies that have investigated self-esteem development using adolescence as its starting point. Furthermore, by considering the extent to which the ways that "moving away from" or "moving against" the world (Caspi et al., 1987, 1988; Horney, 1950) might allow for the development of positive self-views, we provide tentative evidence that shyness and aggressiveness in early childhood may not be key factors for later self-esteem development, adding to the sparse literature on the origins of self-esteem. Below we discuss the implications of our findings with regards to our research aims.

4.1. Rank-order stability of self-esteem from middle childhood to young adulthood

One aim of the present study was to examine the stability in individual differences in self-esteem from middle childhood to young adulthood. These findings are in line with previous research that shows that stability is lower in adolescence (Alsaker & Olweus, 1992; Block & Robins, 1993) than it is in young adulthood (Trzesniewski et al., 2003), but differ markedly from estimates that have been previously found for childhood, in which stability is found to be substantially lower. One factor that may have contributed to this is our use of a first-order autoregressive model to examine rank-order stability in the present study, allowing for estimates that are free of measurement error and that also take into account the dependency of each assessment on a prior assessment.

What else could explain the high levels of stability we found in middle childhood? Research suggests that self-esteem during this time is influenced by self-perceived competence in domains that the individual finds to be important to the individual, as well as social support from peers (Harter, 2012). At the ages of 9 and 10, the individuals in our sample were nearing the end of elementary school. Most German children attend elementary school for four years from the 6 years of age, and it could be that such high levels of stability manifested because the social environment during this time was relatively stable, especially in comparison to the impending transition into junior high as well as pubertal changes to come, two factors that have been demonstrated to influence the stability of self-esteem (Eccles, Midgley, & Adler, 1984; Simmons, Rosenberg, & Rosenberg, 1973).

4.2. Mean-level change in self-esteem from middle childhood to young adulthood

Another aim of the present study was to examine the trajectory of self-esteem through the important developmental transitions between childhood and adolescence, and adolescence and young adulthood. The findings from the present study contribute to the literature by being the first to track self-esteem in the same group of individuals during middle childhood, adolescence, and young adulthood by linking two measures of self-esteem using item response theory methods. Additionally, by examining how self-esteem changes starting from middle childhood, we were able to track self-esteem during a period that has been understudied within the lifespan literature. Our results show that there is a normative, or average trajectory of self-esteem from middle childhood to young adulthood, such that self-esteem is relatively high in

middle childhood, decreases into late childhood, and then steadily increases into young adulthood.

The findings from the present study provide longitudinal support for previous accounts that self-esteem is high in childhood (Harter, 2012; Marsh, Craven, & Debus, 1991; Robins et al., 2002), and decreases into adolescence (Robins et al., 2002). These results indicate that the self-esteem of individuals in our sample was most impacted at 12-years-old, as it showed a mean-level decrease that did not vary much across individuals. Biological (e.g., pubertal), psychological (e.g., increased capacity for abstract thought) and social (e.g., changing schools, shifts in social relationships) changes are likely to be implicated in this decrease.

We also found that self-esteem increased from the point the individuals in our sample were 12-years-old, and our results differ somewhat from previous cross-sectional findings that show that self-esteem continues to decrease throughout the adolescent years, starting to increase only around the time individuals are 18-years-old (Robins et al., 2002). Instead, our findings are more consistent with extant longitudinal work that show increases in self-esteem beginning at 13- and 14-years-old (Orth et al., 2015; von Soest et al., 2016). Our findings also correspond with previous longitudinal research that shows self-esteem increases from adolescence into young adulthood (Hutteman et al., 2015; Orth et al., 2012). Additionally, the effect size estimate from our age 13 assessment to our age 29 assessment (Cohen's $d = 0.29$) is similar to the effect size estimates shown in previous research (Cohen's $d \sim .25$ between age 15 and age 30 in Orth et al., 2015; Cohen's $d = .30$ between age 13 to age 31 in von Soest et al., 2016). The increases that we found in self-esteem from adolescence to young adulthood suggest that many of the individuals in our sample were successful in negotiating the challenges that accompany these important transitions (Hutteman, Hennecke, Orth, Reitz, & Specht, 2014).

4.3. Early childhood shyness and aggressiveness as antecedents of self-esteem development

In using piecewise growth models, we were able to identify a normative trajectory of self-esteem from middle childhood to young adulthood, and also ascertain that there were individual differences in self-esteem levels when children were 9-years old and subsequent changes in self-esteem from age 12 to 29. Given that we were able to model self-esteem development from an age in which self-esteem is thought to take on substantial meaning (Harter, 2012), a secondary aim of the present study was to examine whether individual differences in shyness and aggressiveness in early childhood predicted individual differences in the self-esteem trajectory.

Sociometer theory (Leary et al., 1995) identifies social acceptance as a key contributor to self-esteem, and we reasoned that acting in ways that hinder successful social interactions (i.e., possessing high levels of shyness and/or aggressiveness) might negatively affect self-esteem development. Thus, we examined whether individual differences in shyness and aggressiveness in early childhood predicted individual differences in the self-esteem trajectory. We found that these effects were not significant, indicating that children's shyness and aggressiveness in early childhood did not appear to have an impact on self-esteem levels at 9-years-old, as well as their subsequent development. By taking a multi-informant approach to assessing shyness and aggressiveness, the present study provides a multidimensional assessment of these constructs, moving beyond limitations of previous research that may suffer from shared method variance.

Why is it that these crucial ways of interacting to the world, shown in previous literature to be associated with self-esteem (Donnellan et al., 2005; Kemple et al., 1996; Rosenberg, Schooler, & Schoenbach, 1989; Rubin et al., 2009), do not affect its developmental course? It could be that certain types of shyness and aggressiveness, may not always be detrimental for social acceptance, and ultimately for self-esteem development. Although the current study reveals non-significant effects that should be examined in future research, our results lend support to studies that find that being socially withdrawn and aggressive may not be necessarily consequential for peer relationships in childhood (Ladd & Burgess, 1999; Richardson, Hitti, Mulvey, & Killen, 2014; Vaughn, Vollenweider, Bost, Azria-Evans, & Snider, 2003).

Furthermore, although we examined shyness and aggressiveness as predictors of the self-esteem trajectory because we were interested in the effect of early childhood personality on self-esteem development, it is likely that individual differences in the self-esteem trajectory predict later shyness and aggressiveness. Previous research examining self-esteem as a predictor of aggression shows that low self-esteem in late childhood indeed predicts greater likelihood of aggression in early adolescence (Donnellan et al., 2005), suggesting the viewpoint that self-esteem might be better modeled as a cause than as a consequence of shyness and aggressiveness (Orth & Robins, 2014).

4.4. Limitations and future directions

The present study is not without limitation. First, our sample size ($N = 240$) is relatively small. Yet, our analyses were conducted on data from 9 assessments collected over 25 years, and are modestly powered. Nevertheless, the robustness of findings and generalizability of results should be interpreted with these considerations in mind. Although we do not imagine that it is likely that researchers will conduct direct replications of the current research given its long timespan, we encourage more research to include the examination of self-esteem development from childhood to adolescence, because this is where the data are most lacking, probably due to the complexity of measuring self-esteem adequately through a developmental transition (e.g., in many longitudinal studies, like ours, self-esteem is measured by one measure in childhood and another in adolescence). Recently, a lifespan measure of self-esteem (Harris, Donnellan, & Trzesniewski, 2017) has been introduced that allows for the comparison of the same measure across the lifespan.

Second, it is important to note that we used a linking procedure that assumes that the measures used in our study are assessing the same constructs. Longitudinal studies that examine developmental processes from childhood to adulthood typically employ different measures of the psychological construct of interest to ensure that the assessments are developmentally appropriate. The costs of changing measures make statistical inferences about how the construct of interest develops difficult. In the present study, it was important that our measures change to match the developmental shift that occurred at the cusp of adolescence, resulting in two different measures of each construct being used. For example, the SPPC (Harter, 1985) uses simplified language, explicitly references children, and employs a procedure designed to obtain self-ratings from youth. In contrast, the SDQ-III (Marsh & O'Neill, 1985) references complex thoughts and feelings towards the self, explicitly references the self, and uses a Likert-type response scale, making it appropriate for adolescents and adults. However, because we made use of item response theory methods, we were able to overcome the limitations of previous studies and examine the trajectories of self-esteem from childhood to young adulthood, but under the assumption that the different measures were assessing the same underlying construct.

Third, we did not differentiate types of shyness and aggressiveness in the current study. In another study that examined the LOGIC data, shyness in response to strangers in early childhood was not predictive of self-esteem change in middle childhood,

but shyness in the presence of peers was (Asendorpf & van Aken, 1994). Extant research also suggests that when aggression is differentiated, relational aggression is sometimes predictive of greater social acceptance, while physical aggression is generally predictive of lesser social acceptance (Cillessen & Mayeux, 2004; Nelson, Robinson, & Hart, 2005).

Fourth, we used a piecewise growth model to characterize our data, which led us to specify the childhood slope (i.e., ages 9, 10, and 12) with assessments that were closer together than for the youth slope (i.e., ages 17, 23, and 29). On the one hand, it is possible that a more precise picture of development is obtained when using more assessments during a shorter time span, and that a less precise picture of development is obtained when using fewer assessments during a longer time span. On the other hand, it could be that a longer time span allows for a longitudinal trend to be shown in a less noisy, more reliable manner. However, previous theory and research regarding self-esteem development indicates that normative changes in self-esteem occur the most in childhood and adolescence, and is more uniform into young adulthood. Nevertheless, we encourage future research to include more assessments during adolescence and young adulthood whenever possible.

5. Conclusion

In the present study, we examined the development of self-esteem in a sample of individuals followed longitudinally from middle childhood to young adulthood. We applied item response theory methods to link data from two different self-esteem measures, allowing the present study to track self-esteem longitudinally from a developmental period that has been neglected in the lifespan literature – middle childhood. We found that rank-order stability of self-esteem was high in middle childhood, low in adolescence, and highest in young adulthood. On average, self-esteem levels were relatively high in middle childhood, decreased into adolescence, but increased into young adulthood. Furthermore, shyness and aggressiveness as rated by parents, teachers, and observers in early childhood did not influence self-esteem development.

Acknowledgments

We are grateful to Phil Chalmers, Kevin Grimm and Jonathan Helm for their statistical advice, to Ziyan Luan and Lisanne de Moor for their help with data organization, and to Gabriella Harari, Duygu Taşfiliz and Keith Widaman for useful discussions about this research. The title of this manuscript comes from the songs of Townes van Zandt and Neil Halstead.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jrp.2017.07.001.

References

Alsaker, F. D., & Olweus, D. (1992). Stability of global self-evaluations in early adolescence: A cohort longitudinal study. *Journal of Research on Adolescence, 2*, 123–145. http://dx.doi.org/10.1207/s15327795jra0202_2.

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology, 53*, 27–51. http://dx.doi.org/10.1146/annurev.psych.53.100901.135231.

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.

Asendorpf, J. B. (1990). Development of inhibition during childhood: Evidence for situational specificity and a two-factor model. *Developmental Psychology, 26*, 721–730. http://dx.doi.org/10.1037/0012-1649.26.5.721.

Asendorpf, J. B. (1992). A brunswikean approach to trait continuity: Application to shyness. *Journal of Personality, 60*, 53–77. http://dx.doi.org/10.1111/j.1467-6494.1992.tb00265.x.

Asendorpf, J. B., & van Aken, M. G. (1994). Traits and relationship status: Stranger versus peer group inhibition and test intelligence versus peer group competence as early predictors of later self-esteem. *Child Development, 65*(6), 1786–1798. http://dx.doi.org/10.2307/1131294.

Asendorpf, J. B., & van Aken, M. A. (1999). Resilient, overcontrolled, and undercontrolled personality prototypes in childhood: Replicability, predictive power, and the trait-type issue. *Journal of Personality and Social Psychology, 77*, 815–832. http://dx.doi.org/10.1037/0022-3514.77.4.815.

Asendorpf, J. B., & Van Aken, M. A. G. (2003). Personality–relationship transaction in adolescence: Core versus surface personality characteristics. *Journal of Personality, 71*, 629–666. http://dx.doi.org/10.1111/1467-6494.7104005.

Asendorpf, J. B., Denissen, J. J., & van Aken, M. A. (2008). Inhibited and aggressive preschool children at 23 years of age: Personality and social transitions into adulthood. *Developmental Psychology, 44*, 997–1011. http://dx.doi.org/10.1037/0012-1649.44.4.997.

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: Academic Press.

Bentler, P. M., & Chou, C. (1987). Practical issues in structural equation modeling. *Sociological Methods and Research, 16*, 78–117. http://dx.doi.org/10.1177/0049124187016001004.

Block, J., & Robins, R. W. (1993). A longitudinal study of consistency and change in self-esteem from early adolescence to early adulthood. *Child Development, 64*, 909–923. http://dx.doi.org/10.1111/j.1467-8624.1993.tb02951.x.

Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika, 50*, 229–242.

Boomsma, A. (1982). Robustness of LISREL against small sample sizes in factor analysis models. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirection observation: Causality, structure, prediction (Part 1)* (pp. 149–173). Amsterdam, Netherlands: North Holland.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Caspi, A., & Shiner, R. (2011). Temperament and personality. In M. Rutter, D. Bishop, D. Pine, S. Scott, J. Stevenson, E. Taylor, & A. Thapar (Eds.), *Rutter's child and adolescent psychiatry* (5th ed., pp. 182). Malden, Massachusetts, USA: Blackwell Publishing Limited.

Caspi, A., Elder, G. H., & Bem, D. J. (1987). Moving against the world: Life-course patterns of explosive children. *Developmental Psychology, 23*, 308–313. http://dx.doi.org/10.1037/0012-1649.23.2.308.

Caspi, A., Elder, G. H., & Bem, D. J. (1988). Moving away from the world: Life-course patterns of shy children. *Developmental Psychology, 24*, 824–831. http://dx.doi.org/10.1037/0012-1649.24.6.824.

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software, 48*, 1–29.

Cheek, J. M., & Buss, A. H. (1981). Shyness and sociability. *Journal of Personality and Social Psychology, 41*, 330–339. http://dx.doi.org/10.1037/0022-3514.41.2.330.

Cillessen, A. H., & Mayeux, L. (2004). From censure to reinforcement: Developmental changes in the association between aggression and social status. *Child Development, 75*(1), 147–163. http://dx.doi.org/10.1111/j.1467-8624.2004.00660.x.

Cooley, C. (1902). *Human nature and the social order*.

Coplan, R. J., Prakash, K., O'Neil, K., & Armer, M. (2004). Do you 'want' to play? Distinguishing between conflicted shyness and social disinterest in early childhood. *Developmental Psychology, 40*(2), 244–258. http://dx.doi.org/10.1037/0012-1649.40.2.244.

Coplan, R. J., & Rubin, K. H. (2010). Social withdrawal and shyness in childhood. In K. H. Rubin & R. J. Coplan (Eds.), *The development of shyness and social withdrawal*. New York, NY: Guilford Press.

Crozier, W. (Ed.) (1995). Shyness and self-esteem in middle childhood. *British Journal of Educational Psychology, 65*, 85–95. http://dx.doi.org/10.1111/j.2044-8279.1995.tb01133.x.

Curran, P. J., Hussong, A. M., Cai, L., Huang, W., Chassin, L., Sher, K. J., & Zucker, R. A. (2008). Pooling data from multiple longitudinal studies: The role of item response theory in integrative data analysis. *Developmental Psychology, 44*(2), 365–380. http://dx.doi.org/10.1037/0012-1649.44.2.365.

de Ayala, R. J. (Ed.). (2009). *Theory and practice of item response theory (methodology in the social sciences)*. New York, NY: Guilford Press.

Denissen, J. J. A., Asendorpf, J. B., & Van Aken, M. A. G. (2008). Childhood personality predicts long-term trajectories of shyness and aggressiveness in the context of demographic transitions in emerging adulthood. *Journal of Personality, 76*, 67–99. http://dx.doi.org/10.111/j.1467-6494.2007.00480.x.

Denissen, J. J. A., Thomaes, S., & Bushman, B. J. (2017). Self-regulation and aggression: Aggression-provoking cues, individual differences, and self-control strategies (in press).

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*(6), 634–645. http://dx.doi.org/10.1016/j.jrp.2012.07.005.

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.

Eccles, J. S., Midgley, C., & Adler, T. (1984). Grade-related changes in the school environment. *Advances in Motivation and Achievement, 3*, 238–331.

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*(3), 830–847. http://dx.doi.org/10.2307/1131221.

Embretson, S., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: L. Erlbaum Associates.

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.

Gest, S. D. (1997). Behavioral inhibition: Stability and associations with adaptation from childhood to early adulthood. *Journal of Personality and Social Psychology, 72*(2), 467–475. http://dx.doi.org/10.1037/0022-3514.72.2.467.

Göttert, R., & Asendorpf, J. (1989). Eine deutsche Version des California- Child-Q-Sort, Kurzform [A German short version of the California Child Q-set]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 21*, 70–82.

Harris, M. A., Donnellan, M. B., & Trzesniewski, K. H. (2017). Lifespan self-esteem scale: Initial validation of a new measure of global self-esteem. *Journal of Personality Assessment*.

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*(4), 995–1013. http://dx.doi.org/10.1111/cdev.12356.

Hart, D., Hofmann, V., Edelstein, W., & Keller, M. (1997). The relation of childhood personality types to adolescent behavior and development: A longitudinal study of Icelandic children. *Developmental Psychology, 33*(2), 195. http://dx.doi.org/10.1037/0012-1649.33.2.195.

Harter, S. (Ed.). (1985). *The self-perception profile for children: Revision of the perceived competence scale for children*. Denver, CO: University of Denver.

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

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.

Horney, K. (1950). *Neurosis and human growth*. New York: Norton.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. http://dx.doi.org/10.1080/10705519909540118.

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.

Hutteman, R., Denissen, J., Asendorpf, J., & Van Aken, M. (2009). Changing dynamics in problematic personality: A multiwave longitudinal study of the relationship between shyness and aggressiveness from childhood to early adulthood. *Development and Psychopathology, 21*, 1083–1094. http://dx.doi.org/10.1017/S0954579409990058.

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*(3), 267–278.

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.

Kemple, K. M., David, G. M., & Wang, Y. (1996). Preschoolers' creativity, shyness, and self-esteem. *Creativity Research Journal, 9*, 317–326. http://dx.doi.org/10.1207/s15326934crj0904_3.

Kerr, M. (2000). Childhood and adolescent shyness in long-term perspective: Does it matter? In W. Ray Crozier (Ed.), *Shyness: Development, consolidation and change* (pp. 64–87). Routledge: NY.

Kjeldsen, A., Nilsen, W., Gustavson, K., Skipstein, A., Melkevik, O., & Karevold, E. B. (2016). Predicting well-being and internalizing symptoms in late adolescence from trajectories of externalizing behavior starting in infancy. *Journal of Research on Adolescence*. http://dx.doi.org/10.1111/jora.12252.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.

Kolen, M. J., & Brennan, R. L. (Eds.). (2004). *Test equating, scaling, and linking*. New York, NY: Springer.

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.

Ladd, G. W., & Burgess, K. B. (1999). Charting the relationship trajectories of aggressive, withdrawn, and aggressive/withdrawn children during early grade school. *Child Development, 70*(4), 910–929. http://dx.doi.org/10.1111/1467-8624.00066.

Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as interpersonal monitor: The sociometer hypothesis. *Journal of Personality and Social Psychology, 68*, 518–530. http://dx.doi.org/10.1037/0022-3514.68.3.518.

Luan, Z., Poorthuis, A. G., Hutteman, R. H., Asendorpf, J. B., Denissen, J. J. A., & van Aken, M. A. G. (2017). See me through my eyes: Adolescent-parent agreement in personality predicts later self-esteem development. *International Journal of Behavioral Development*. http://dx.doi.org/10.1177/0165025417690263.

Marsh, H. W., Craven, R. G., & Debus, R. (1991). Self-concepts of young children 5 to 8 years of age: Measurement and multidimensional structure. *Journal of Educational Psychology, 83*, 377–392. http://dx.doi.org/10.1037/0022-0663.83.3.377.

Marsh, H. W., & O'Neill, R. (1985). Self-Description Questionnaire III (SDQ III): The construct validity of multidimensional self-concept ratings by adolescents. *Journal of Educational Measurement, 21*, 153–174.

Mead, G. (1934). *Mind, self, and society from the standpoint of a social behaviorist*. Chicago: University of Chicago Press.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nelson, D. A., Robinson, C. C., & Hart, C. H. (2005). Relational and physical aggression of preschool-age children: Peer status linkages across informants. *Early Education & Development, 16*(2), 115–140. http://dx.doi.org/10.1207/s15566935eed1602_2.

Orth, U. (2017). 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*. http://dx.doi.org/10.1037/pspp0000143.

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. (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., & 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., 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.

Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit (Language and personality structure: Validity of the five-factor model of personality)*. Regensburg, Germany: S. Roderer.

R Core Team (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Revelle, W. (2015). *Psych: Procedures for personality and psychological research*. Evanston, Illinois, USA: Northwestern University.

Richardson, C. B., Hitti, A., Mulvey, K. L., & Killen, M. (2014). Social exclusion: The interplay of group goals and individual characteristics. *Journal of Youth and Adolescence, 43*(8), 1281–1294. http://dx.doi.org/10.1007/s10964-013-9967-8.

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., 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.

Robins, R. W., Donnellan, M. B., Widaman, K. F., & Conger, R. D. (2010). Evaluating the link between self-esteem and temperament in Mexican origin early adolescents. *Journal of Adolescence, 33*, 403–410. http://dx.doi.org/10.1016/j.adolescence.2009.07.009.

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.

Rodriguez, D., Wigfield, A., & Eccles, J. S. (2003). Changing competence perceptions, changing values: Implications for youth sport. *Journal of Applied Sport Psychology, 15*, 67–81. http://dx.doi.org/10.1080/10413200305403.

Rosenberg, M., Schooler, C., & Schoenbach, C. (1989). Self-esteem and adolescent problems: Modeling reciprocal effects. *American Sociological Review*, 1004–1018.

Rothbart, M. K. (2011). *Becoming who we are*. New York, NY: Guilford Press.

Rubin, K. H., Chen, X., McDougall, P., Bowker, A., & McKinnon, J. (1995). The Waterloo Longitudinal Project: Predicting internalizing and externalizing problems in adolescence. *Development and Psychopathology, 7*, 751–764. http://dx.doi.org/10.1017/S0954579400006829.

Rubin, K. H., Coplan, R. J., & Bowker, J. C. (2009). Social withdrawal in childhood. *Annual Review of Psychology, 60*, 141–171. http://dx.doi.org/10.1146/annurev.psych.60.110707.163642.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement, 34*(4), 100.

Schinka, K. C., van Dulmen, M. M., Mata, A. D., Bossarte, R., & Swahn, M. (2013). Psychosocial predictors and outcomes of loneliness trajectories from childhood to early adolescence. *Journal of Adolescence, 36*(6), 1251–1260. http://dx.doi.org/10.1016/j.adolescence.2013.08.002.

Schneider, W., & Bullock, M. (Eds.). (2009). *Human development from early childhood to early adulthood: Findings from a 20 year longitudinal study*. New York, NY: Psychology Press.

Simmons, R. G., Rosenberg, F., & Rosenberg, M. (1973). Disturbance in the self-image at adolescence. *American Sociological Review*, 553–568.

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.

Sturaro, C., Denissen, J. J., van Aken, M. A., & Asendorpf, J. B. (2008). Person-environment transactions during emerging adulthood. *European Psychologist, 13*, 1–11. http://dx.doi.org/10.1027/1016-9040.13.1.1.

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.

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.

Vaughn, B. E., Vollenweider, M., Bost, K. K., Azria-Evans, M. R., & Snider, J. B. (2003). Negative interactions and social competence for preschool children in two samples: Reconsidering the interpretation of aggressive behavior for young children. *Merrill- Palmer Quarterly, 49*, 245–278.

von Soest, T., Wichstrøm, 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.

Wegener, W. (1988). *Kritik des Prestiges [Critique of prestige]*. Opladen, Germany: Westdeutscher Verlag.

Weinert, F. E., & Schneider, W. (Eds.). (1999). *Individual development from 3 to 12: Findings from the Munich longitudinal study*. Cambridge, UK: Cambridge University Press.

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.