Self-Concept of Students in Childhood and Adolescence:

Structural and Developmental Issues

Dipl.-Psych. Isabelle Schmidt

Dissertation submitted to the

University of Trier, Department of Psychology

(Faculty I – Department of Psychology

Giftedness Research and Education), Trier

To obtain the degree of doctor rerum naturalium (Dr. rer. nat.) in Psychology

Submitted: May 2, 2018

Supervisors:

Prof. Dr. Franzis Preckel

University of Trier, Chair of Giftedness Research and Education

Prof. Dr. Martin Brunner

University of Potsdam, Chair for Quantitative Methods in Educational Sciences

Place of Dissertation: Trier, Germany

Table of Contents

Acknowledgement

Foremost, I would like to express my sincere gratitude to my advisors Prof. Dr. Franzis Preckel and Prof. Dr. Martin Brunner for their support, patience, motivation, enthusiasm, and immense knowledge they provided me. Their guidance inspired me throughout the entire process of researching and writing this thesis.

Further, I would like to thank my co-workers for their support in several ways.

I would like to thank my family for giving birth to me at the first place and supporting me throughout my life.

Abstract

Fostering positive and realistic self-concepts of individuals is a major goal in education worldwide (Trautwein & Möller, 2016). Individuals spend most of their childhood and adolescence in school. Thus, schools are important contexts for individuals to develop positive self-perceptions such as self-concepts. In order to enhance positive self-concepts in educational settings and in general, it is indispensable to have a comprehensive knowledge about the development and structure of self-concepts and their determinants. To date, extensive empirical and theoretical work on antecedents and change processes of self-concept has been conducted. However, several research gaps still exist, and several of these are the focus of the present dissertation. Specifically, these research gaps encompass (a) the development of multiple self-concepts from multiple perspectives regarding stability and change, (b) the direction of longitudinal interplay between self-concept facets over the entire time period from childhood to late adolescence, and (c) the evidence that a recently developed structural model of academic self-concept (nested Marsh/Shavelson model [Brunner et al., 2010]) fits the data in elementary school students, (d) the investigation of structural changes in academic self-concept profile formation within this model, (e) the investigation of dimensional comparison processes as determinants of academic self-concept profile formation in elementary school students within the internal/external frame of reference model (I/E model; Marsh, 1986), (f) the test of moderating variables for dimensional comparison processes in elementary school, (g) the test of the key assumptions of the I/E model that effects of dimensional comparisons depend to a large degree on the existence of achievement differences between subjects, and (h) the generalizability of the findings regarding the I/E model over different statistical analytic methods.

Thus, the aim of the present dissertation is to contribute to close these gaps with three studies. Thereby, data from German students enrolled in elementary school to secondary school education were gathered in three projects comprising the developmental time span from childhood to adolescence (ages 6 to 20). Three vital self-concept areas in childhood and adolescence were investigated: general self-concept (i.e., self-esteem), academic self-concepts (general, math, reading, writing, native language), and social self-concepts (of acceptance and assertion). In all studies, data were analyzed within a latent variable framework.

In Study 1, the development of self-esteem, general academic self-concept, and social self-concept of acceptance and assertion was investigated from late childhood to late adolescence (secondary school students from grade 5 to grade11; *N* = 1,163) from multiple perspectives, that is, on a sample level (mean-level change and rank-order stability) and individual level (interindividual differences in intraindividual change). Further, the causal flow between self-concepts facets was investigated. Results indicated that similarities and differences in the development of self-concept exist across the four self-concepts. All self-concept facets became more stable in rank-order over time and were comparably stable. Results from second order latent growth curve models indicated different developments. Academic self-concept declined in mean level, self-esteem stayed constant, and both social self-concepts increased in mean level. Further, shapes of mean-level change differed between facets. Academic self-concept and social self-concept of assertion showed a non-linear mean-level change, and all other self-concept facets showed a linear mean-level change. In all self-concepts, interindividual differences in intraindividual change were found. Results from cross-lagged panel analysis showed that the pattern of the causal flow between self-concept facets was relatively balanced and thus not providing evidence for a predominance of the causal flow.

In Study 2, the recently developed structural model of academic self-concept, the nested Marsh/Shavelson model of academic self-concept (Brunner et al., 2010), was tested in samples of elementary school students (grade 1 to grade 4; Total *N* = 3,779) using confirmatory factor analysis. Moreover, structural changes in academic self-concept formation (i.e., differentiation and integration) was tested. Further, the effects of dimensional comparisons on the development of subject-specific academic self-concepts, which are expected to be a determinant of multidimensionality of academic self-concepts as posited in the I/E model. Thereby, the nested Marsh/Shavelson model of academic self-concept and a first-order correlated factor model was used. The assumptions of the I/E model were examined for the native language self-concept, which comprises the two related core skill academic self-concepts of reading and writing, thus going beyond the majority of previous studies which investigated only reading academic self-concept. Moreover, a moderator for dimensional comparison processes, that is, the size of grade difference was tested. Results showed that the nested Marsh/Shavelson model provided a good fit to the data. Structural changes, that is, differentiation and integration took place early in elementary school during the first two school years for both structural models of academic self-concept (nested Marsh/Shavelson model and first-order correlated factor model). Further, the effect of dimensional comparison processes as posited in the I/E model could be shown in grades 3 and 4 across different structural models. Finally, the size of the grade difference moderated the effect of dimensional comparison processes, that is, with increasing grade difference the size of dimensional comparisons increases.

At last, in Study 3 the key assumptions of the I/E model, that effects of dimensional comparisons depend to a large degree on the existence of achievement differences between subjects, was investigated in eight samples of elementary and secondary school students (grades 3 to 8; *N*s = 326–878). Further, the generalizability of findings over different grades and self-concept scales was tested in a manifest and latent variable framework with a seldom-used analytic approach. Results indicated that the effect of dimensional comparison processes, as posited in the I/E model, generalized across the analytic method, grade levels, and different self-concept scales.

Findings are discussed with respect to the research aims of acquiring more comprehensive knowledge on the structure and development of significant self-concept in childhood and adolescence and their determinants. In addition, theoretical and practical implications derived from the findings of the present studies are outlined. Strengths and limitations of the present dissertation are discussed. Finally, an outlook for future research on self-concepts is given.

Index of publication

The present dissertation is structured into six chapters. In Chapter 1, the rationale for the thesis is outlined. In Chapter 2, the theoretical background of the three studies is presented. Chapter 3 contains an original research article, which is prepared for submission to a journal. Chapter 4 and Chapter 5 include original research articles, which have been published in peer-reviewed journals.

In Chapter 6, the results of the three articles are summarized and reflected upon with respect to their aims and contribution to theory and practice. The author of the present thesis is the first author of all three articles. All publications are presented in the published or prepared form except for minor changes particularly regarding format, layout, and citation style. All articles were prepared with support from co-authors. The following paragraph lists the co-authors of each article.

Study 1	(Chapter 3)	Schmidt, I., Brunner, & Preckel, F. Developmental dynamics of adolescent students' self-esteem, academic self-concept and social self-concepts: Results from a six-year longitudinal study. (Manuscript prepared for submission)
Study 2	(Chapter 4)	Schmidt, I., Brunner, M., Keller, L., Scherrer, V., Wollschläger, R., Baudson, T. G., & Preckel, F. (2017). Profile formation of academic self-concept in elementary school students in grades 1 to 4. *PloS one, 12*(5), e0177854. https://doi.org/10.1371/journal.pone.0177854
Study 3	(Chapter 5)	Schmidt, I., Brunner, M., & Preckel, F. (2017). Effects of achievement differences for internal/ external frame of reference model investigations: A test of robustness of findings over diverse student samples. *British Journal of Educational Psychology.* Advance online publication. https://doi:org/10.1111/bjep.12198

1 Introduction

"Positive self-esteem is clearly a psychological commodity, a resource that is important for us to foster in our children and adolescents if we want them to lead productive and happy lives."

(Harter, 2005, p. 977).

The term *self* is generally used in reference to the conscious reflection of one's own being or identity, as an object separate from others or from the environment (Huitt, 2011, p.1). There are several ways to think about the self (see Filipp & Mayer, 2005) with *self-concept* as one of the most prominent ways in research and practice. See in this respect the distinction between *I-self (the self as a knower)* and *Me-self (the self as known)* by William James (1890, 1892). He characterized the Me-self as an "empirical aggregate of things objectively known" (p. 197). The *Me-self* is notably reflected in the conscious construction of one's self-concept.

Self-concepts are important variables that foster positive outcomes in different areas of life like social relations, academics, and mental and physical health (e.g., Baumeister, Campbell, Krueger, & Vohs, 2003; Craven & Marsh, 2008; Trzesniewski et al., 2006). In this regard, Craven and Marsh (2008) remarked that: "Hence, a positive self-belief is valued as a 'hot' variable that makes good things happen, and is fundamental to the realization of full human potential in a range of settings" (p. 107).

The development of the self is a lifelong process (Erikson, 1971), but childhood and adolescence are important periods for the development of a positive self-concept. In particular, the investigation of the development of self-concepts in childhood and adolescence is of importance because a positive development of self-concepts during this time has long-lasting positive implications for academic career, behavioral adjustment, mental health, and social relations (e.g., Baumeister et al., 2003; Fergusson & Woodward, 2002; Steiger, Allemand, Robins, & Fend, 2014; Trzesniewski et al., 2006). Childhood and adolescence are periods in which many physical, psychological, cognitive, and context transitions occur and in which individuals experience several challenges in different areas of life (Harter, 2012b). However, it is challenging to develop positive self-concepts and maintain positive self-concepts during changing contexts and cognitive and physical changes (e.g., puberty).

As children and adolescents spend most of their times in schools (Eccles & Roeser, 2009), educational contexts and educational practice have a major impact on individuals’ self-concepts (e.g., Aldrup, Klusmann, Lüdtke, Göllner, & Trautwein, 2018; Burnett, 2003; Eccles et al., 1989; Eccles & Roeser, 2009). In educational psychology, self-concepts, particularly academic self-concepts, are targeted because they facilitate a wide range of academic behavior and outcome (Marsh, Martin, Yeung, & Craven, 2017) (see Subsection 2.1.2). In accordance to the OECD, school engagement including self-concepts are “closely tied to students’ economic success and long-term health and wellbeing” (OECD, 2003, p. 9). Therefore, it is widely accepted that fostering positive but realistic self-concepts is a major educational goal worldwide (Marsh & Hau, 2003b; Trautwein & Möller, 2016).

To meet this challenge, researchers investigate the structure and the development of self-concepts, as well as their determinants. Thus, it is the overall goal of the present dissertation to contribute, in three studies, to existing knowledge.

Up to now, many researchers have investigated the structure and development of self-concept. In the 1970s, Shavelson et al. (1976) laid an important foundation for self-concept research. Reviewing previous research, they proposed a theoretical model of self-concept, making assumptions about the structure and the development of self-concept. They posited a hierarchical and multifaceted structured self-concept. The work by Shavelson et al. has widely influenced subsequent self-concept research, particularly in educational psychology (e.g., Marsh et al., 2017). Nowadays, it is widely accepted across different psychological disciplines that the self-concept is a multifaceted, structured construct consisting of more general and more specific facets (e.g., Baumeister et al., 2003; Kernis, 2006; Marsh et al., 2017). In the present dissertation, three vital self-concept facets are targeted: general self-concept (i.e., self-esteem), and the specific self-concepts (academic self-concepts [general, math, reading, writing, native language]) and social self-concepts (self-concept of acceptance and assertion).

Despite the acceptance that self-concept is a multifaceted construct, how to model the hierarchical aspect of self-concept is a matter of debate (see, e.g., Arens & Morin, 2016; Trautwein & Möller, 2016). Empirical work testing the proposed structural model of Shavelson et al. led to refinements and revisions of the model that was originally proposed. For instance, in contrast to the hierarchical structure by Shavelson et al. (1976), Marsh and colleagues (see, e.g., Marsh, 1990e) found that there is one general academic self-concept is at the apex and that this academic self-concept is multidimensional in the sense that individuals think of themselves as a verbal or math person (Marsh & Hau, 2004). This led to the Marsh/Shavelson model of academic self-concept with two (instead of one) uncorrelated general academic self-concepts at the apex—math domain general and verbal domain general (a detailed description is given in Section 2.1).

However, some research findings throw doubt on this model so that a new structural model, the nested Marsh/Shavelson model (Brunner et al., 2010), was recently proposed. In this model, the hierarchy assumption of academic self-concept by Shavelson et al. and the multidimensionality of academic self-concept are combined in a specific kind of factorial model (as described in Section 2.1). Most importantly, this model was statistically better able to describe the structure in comparison to the Marsh/Shavelson model of academic self-concept in confirmatory factor analyses (see Brunner et al., 2010). However, although studies have provided empirical evidence for secondary school students, no study has investigated this model in elementary school students. Thus, it is an explicit aim of the present dissertation to test the nested Marsh/Shavelson model of academic self-concept in elementary school students.

To explain the multidimensionality of academic self-concept (i.e., individuals think of themselves as a verbal or math person), Marsh (1986) developed the I/E model in which dimensional comparison processes are considered as a major determinant of subject-specific academic self-concept and academic self-concept formation (as described in Section 2.4). This led to the focus of dimensional comparison processes as a major determinant of academic self-concepts (see Möller, 2016). Thus, the I/E model, with its assumptions of dimensional comparison processes and social comparison processes as determinants of subject-specific academic self-concept and academic self-concept formation, is also in the focus of the present dissertation.

Moreover, the influence of dimensional comparison processes for the development of subject-specific academic self-concept as posited in the (I/E model; Marsh, 1986) is especially less studied in elementary school students compared to secondary school students. Furthermore, moderators affecting the influence of dimensional comparison processes on subject-specific academic self-concepts have been rarely investigated. Thus, not only the I/E model alone, but also moderating effects on dimensional comparison processes in elementary school are targeted in the present dissertation. Moreover, the I/E model can be investigated with different statistical approaches (as described in Subsection 2.4.1) from which a seldom-used approach is employed in the present dissertation to test the I/E model assumptions. This approach directly addresses the key assumption of the I/E model, that is, that dimensional comparison processes depend on achievement differences. Thus, it is a further aim to investigate the I/E model with this approach.

Besides investigating determinants of academic self-concept formation within the I/E model, structural changes of academic self-concept and the development of the single self-concepts facets are in the focus of research. Several self-concept researchers (Harter, 2012b; Marsh & Ayotte, 2003; Shavelson, Hubner, & Stanton, 1976; Wigfield & Eccles, 2002) assume that self-concepts change with human development. This also applies to the structure of the self-concept changes with age, that is, the relation between self-concept facets change with age (see Section 2.2).

Structural changes can be investigated in different factor models of academic self-concept. An explanation is provided describing why the nested Marsh/Shavelson model can be seen as a suitable approach to investigate structural changes in academic self-concept formation (as described in Subsection 2.2.1). Thus, a further aim of the present dissertation is to investigate structural changes in the nested Marsh/Shavelson model during the elementary school timeframe, which has not been tested to date. Different theoretical assumptions of self-concept development and structural change can be derived from different theories (a detailed description is given in Section 2.4).

Stability and change in self-concept development can be considered from different perspectives, each highlighting different aspects of change: sample-level change (mean-level change, rank-order stability) and individual-level change (interindividual differences in intraindividual change, ipsative stability) (as described in detail in Subsection 2.4.1). Up to now, self-concept researchers have often investigated the development of multiple self-concept facets mainly on a sample level (mean-level change and rank-order stability), but comparably seldom have they investigated the development of multiple self-concept facets on the individual level. Thus, a further aim of the present dissertation is to contribute to the knowledge by investigating multiple self-concept facets (self-esteem, general academic self-concept, social self-concept of acceptance and assertion) on sample level and individual level for the time span from late childhood to late adolescence.

Besides how the single self-concept facets develop, it is a debate in self-concept research how and if self-concept facets are causally related with each other across time (see e.g., Wagner & Valtin, 2004). Up to now, empirical research on that topic is insufficient and inconsistent, and often target only small time spans (e.g., Marsh & Yeung, 1998). Thus, lastly, an explicit aim of the present dissertation is to investigate the longitudinal relations between self-esteem, academic self-concept, and social self-concept of acceptance and assertion from late childhood to late adolescence to shed light on the pattern of the causal flow in this important time span for a positive development of self-concepts.

2 Theoretical background

This chapter addresses the theoretical background on which the three studies are based. To this end, theoretical and empirical work on the structure and development of different self-concept facets and determinants of those in focus in the present dissertation are outlined. Thereby, existing research gaps targeted in the present dissertation are uncovered and described. In doing so, first, the seminal work by Shavelson et al. (1976) on the structure of self-concept is described as well as the substantial structural revisions of the self-concept model, the Marsh/Shavelson model and the nested Marsh/Shavelson model (see Section 2.1). These are targeted in Study 2. Thereafter, the three self-concept areas investigated in the present dissertation (self-esteem, academic self-concepts, social self-concepts) are presented, and findings from previous research on their significance in childhood and adolescence in the educational context and for other areas of life and the short-term and long-term of the development in this period regarding outcomes are reported (see Subsections 2.1.1-2.1.3). In Section 2.2, theoretical assumptions about the development of self-concept are outlined according to the work of Shavelson et al. (1976), Susan Harter, and Herbert Marsh and colleagues. Next, multiple perspectives of stability and change in self-concept development are described (see Section 2.3), which are more or less the basis for all studies in the present dissertation. In Section 2.3, the theoretical positions on the causal flow between self-concept facets and the status of previous empirical research on this topic are presented, which is the topic of investigation in Study 1. Finally, a model that posits dimensional comparisons as a major determinant of academic self-concept formation, the I/E model, moderating variables, and different statistical approaches to investigate the I/E model are outlined in Section 2.4. The I/E model is put to the test in Studies 2 and 3.

2.1 Structure of the self

Until the 1980s self-concept research was characterized by a lack of theoretical basis and a poor quality of self-concept measurement instruments (see e.g., Hattie, 1992; Marsh et al., 2017; Marsh & Hattie, 1996). After reviewing previous self-concept research, Shavelson et al. (1976) proposed a theoretical model of self-concept that caused quite a stir and remarkably stimulated subsequent research on the structure of self-concept.

Shavelson, Hubner, and Stanton (1976, p. 411) referring back to William James (1890, 1892), defined self-concept as "a person's perception of himself. These perceptions are formed through his experience with his environment, […], and are influenced especially by environmental reinforcements and significant others."

According to Shavelson et al., the construct self-concept can be described with seven features:

> (1) It is organized or structured in that people categorize the vast information they have about themselves and relate these categories to one another. (2) It is multifaceted and the particular facets reflect the category system adopted by a particular individual and/or shared by a group. (3) It is hierarchical with perceptions of behavior at the base moving to inferences about self in subareas (e.g., academic—English, history), academic and nonacademic areas, and then to general self-concept. (4) General self-concept is stable but, as one descends the hierarchy, self-concept becomes increasingly situation-specific and, as a consequence, less stable. (5) Self-concept becomes increasingly multi-faceted as the individual develops from infancy to adulthood. (6) It has both a descriptive and an evaluative dimension such that individuals may describe themselves (I am happy) and evaluate themselves (e.g., I do well in school). And (7) it can be differentiated from other constructs such as academic achievement (Shavelson & Bolus, 1982, p. 1).

Feature 7 is described in more detail in Subsection 2.2.2. Features 4 and 5 are focused in Section 2.4.

Regarding the structure of the self-concept (features 2 and 3), Shavelson et al. (1976) assume that the self has a hierarchical and multidimensional structure, with general perceptions of self as a person (i.e., general self-concept, also called self-esteem) at the apex, thus, at the top of the hierarchy (see Figure 1). General self-concept is further divided into an academic and three nonacademic (i.e., physical, social, and emotional) self-concepts. These facets are further divided into more specific components on the next level of hierarchy. For instance, the general academic self-concept factor is subdivided into subject-specific academic self-concepts (e.g., math self-concept, native language self-concept, physics self-concept). Lastly, the evaluation of behavior in specific situations is located at the bottom of the model. Thus, the structure becomes increasingly differentiated when moving from the top to the bottom of the hierarchy.

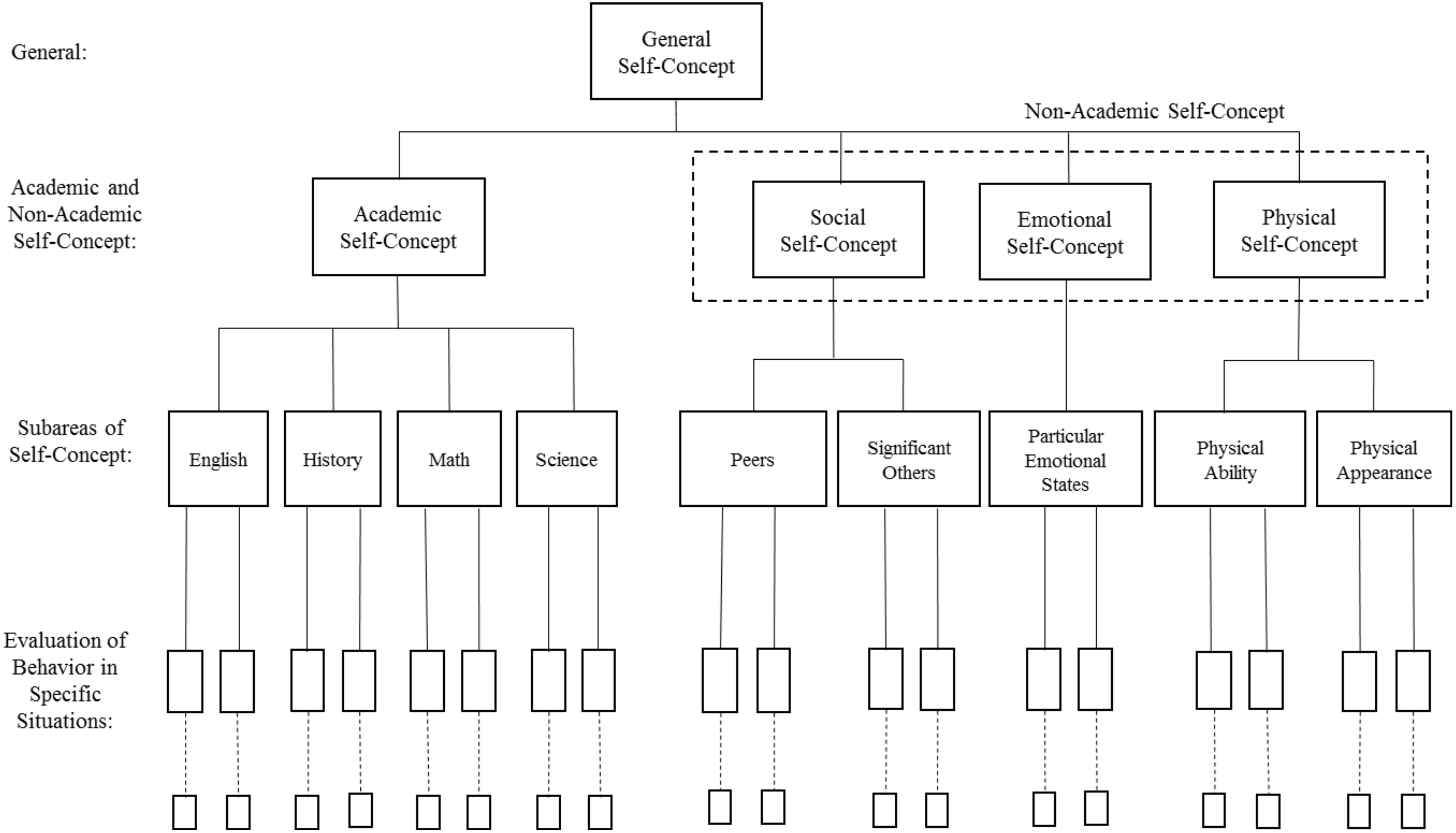


Figure 1. The Shavelson model of self-concept (Shavelson et al., 1976, p. 413)

The hierarchical assumption of self-concept implies that more specific self-concept facets at the base of the hierarchy should correlate more highly with actual behavior than the less specific self-concept facets at the apex of the model. Furthermore, the hierarchical structure suggests that the strength of correlations between self-concept facets varies in a systematic pattern such that general self-concept correlates highest with (general) academic self-concept and next highest with subject-specific self-concepts.

Multidimensionality implies that self-concept facets are related but can be measured as separate constructs. For instance, with reference to the academic section of the model, although general self-concept is expected to correlate with academic self-concept, academic self-concept with subject-specific self-concepts, and general self-concept with subject-specific self-concepts, each of these facets can be interpreted separately (Byrne & Shavelson, 1996).

Furthermore, self-concepts can have an evaluative and a descriptive component. For example, academic self-concepts can have both an evaluative (e.g., “I like math”) and a descriptive component (e.g., “I am good in math”). Both components are strongly related (see e.g., Marsh & Ayotte, 2003), but the descriptive component is more strongly related to academic achievement than the evaluative component (Arens, Yeung, Craven, & Hasselhorn, 2013).

Several studies (e.g., Craven & Marsh, 2008; Marsh & Hattie, 1996) indicate that the global self-concept factor at the apex in the Shavelson et al. model is not distinguished from self-esteem measured with a questionnaire aiming to assess self-esteem as a unidimensional construct like the Rosenberg (Rosenberg, 1965) scale. In line with this, in the present dissertation, the most general self-concept is defined as self-esteem as assessed using items from the Rosenberg scale. Self-esteem is defined as “the individual’s positive or negative attitude toward the self as a totality” (Rosenberg, Schooler, Schoenbach, & Rosenberg, 1995, p. 141).

On the basis of the Shavelson et al. model, Marsh and colleagues developed the Self Description Questionnaire (SDQ)[1] series for different age ranges in order to validate the postulated multidimensional and hierarchical structure of self-concept (see Marsh & Shavelson, 1985). A large body of research with the SDQs (Byrne, 1996; see the reviews by Byrne, 1996; Marsh & Craven, 1997; Marsh & Hattie, 1996) and other multidimensional self-concept measurements (e.g., Self-Perception Profile for Children and for Adolescents (SPP-A and SPP-C, respectively; see Harter, 2012b)[2] supported that self-concept is a multifaceted construct across different age groups, gender, and cultures (Arens, Bodkin-Andrews, Craven, & Yeung, 2014; Esnaola et al., 2018).

Despite the multidimensional aspect of self-concept, regarding the hierarchical aspect of self-concept previous research revealed that hierarchy might be more complicated than expected by Shavelson et al.. Several studies verify the existence of differentiable math and native language self-concepts[3] that are uncorrelated or only weakly correlated, thus challenging the presumption of one global factor representing general academic self-concept (see e.g., Marsh, 1990e; Marsh, Byrne, & Shavelson, 1988).

This finding led to the development of the Marsh/Shavelson model of academic self-concept (Marsh et al., 1988; Marsh, 1990e; Marsh & Shavelson, 1985). In the Marsh/Shavelson model (see Figure 2 a), academic self-concept is no longer considered to be a hierarchical construct with one global factor at the apex but instead a hierarchical construct with two independent higher-order factors at the apex—a verbal and math domain general academic self-concept—were posited. Subject-specific self-concepts are subordinate and can be assigned along a continuum from most

[1] Self Description Questionnaires are available in several versions (cf., Byrne, 1996; see also Arens, Trautwein, and Hasselhorn, 2011; Esnaola, Elosua, and Freeman, 2018).

[2] Self Perception Profiles are available in seven versions comprising the age range from 8 to 60+ years (see Harter, 2012b).

[3] Note, that in the SDQs, native language self-concept is assessed as reading self-concept.

verbal-like (native language and foreign languages) at the one end and more math-like (math, physics, chemistry) at the other end. Math and native language are positioned at the end-points as they are correlated to the smallest degree. Moreover, general academic self-concept is subordinate and in the center of the continuum and influenced simultaneously by the two higher-order academic self-concept factors. Marsh (1990e) developed the Academic Self Description Questionnaire (ASDQ I, II, III)[4] to assess a more diverse variety of academic self-concepts (i.e., school subjects) than has been considered in the pre-existing SDQs. Several studies with the ASDQs have demonstrated some evidence supporting the Marsh/Shavelson model of academic self-concept but also some critical results. For instance, the correlation between verbal and math domain general factors was substantial (r = .50-.70) (Marsh, 1990e). Moreover, foreign language and native language were found to be only modestly correlated, challenging the assumption that they belong to the same higher-order verbal domain factor (Gogol, Brunner, Martin, Preckel, & Goetz, 2017; Marsh, Kong, & Hau, 2001). Further, several studies identified positive correlations of general academic self-concept with both native language and math academic self-concept (e.g., Marsh, 1986; Marsh et al., 1988). These critical findings indicate that native language and math academic self-concept share a large proportion of common variance that may be attributable to general academic self-concept.

[4] The ASDQ-I is intended for preadolescents, the ASDQ-II is intended for adolescents, and the ASDQ-III is intended for late adolescents. The ASDQ items tap into self-concepts in multiple academic areas as well as a student’s overall self-concept (see Byrne, 1996)

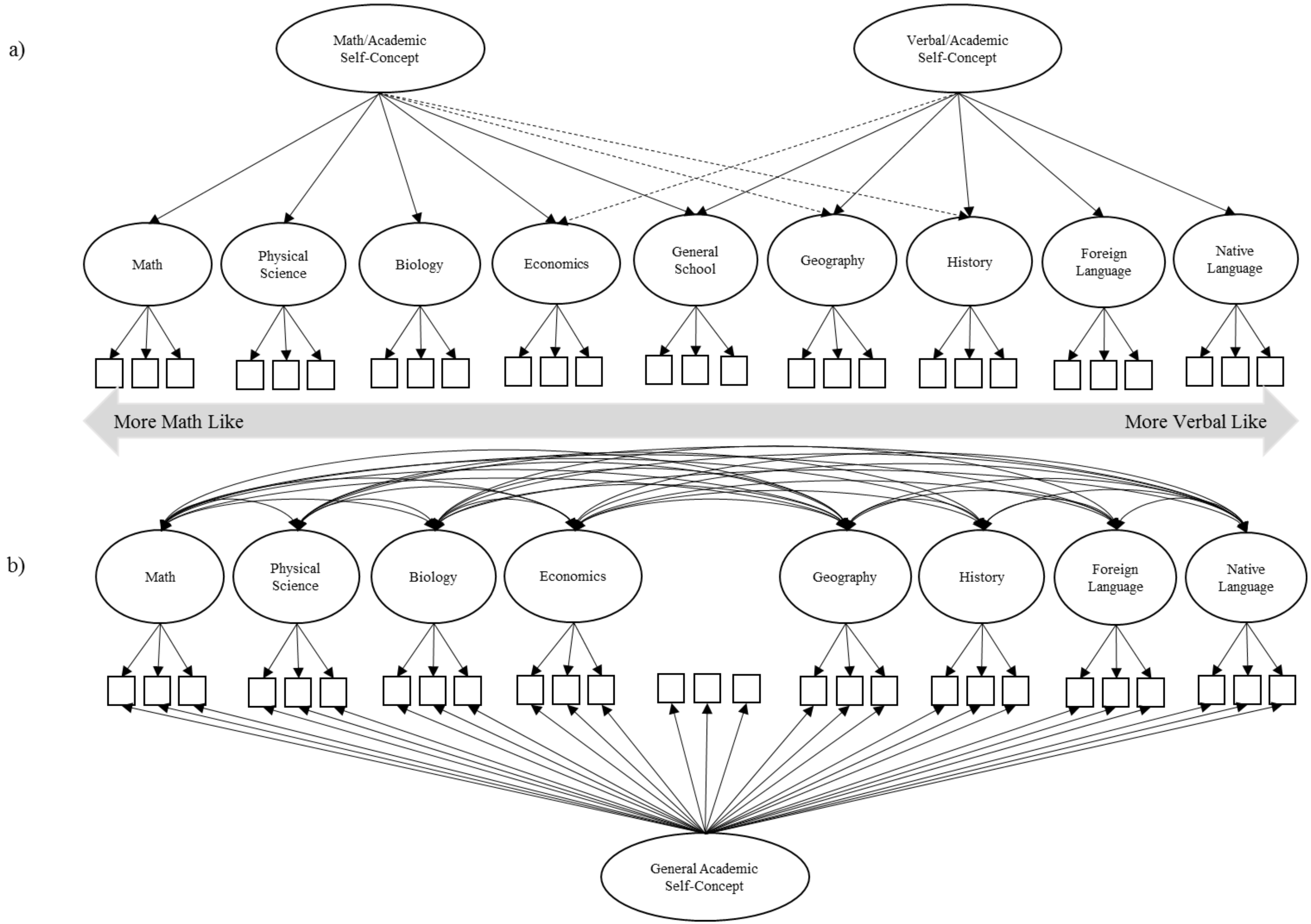


Figure 2. Structural models of academic self-concept. (a) The Marsh/Shavelson model (Marsh et al., 2017, p. 88), (b) an elaboration of the nested Marsh/Shavelson model (Brunner et al., 2010)

Thus, drawing on critical results in self-concept research and recently developed statistical methods in structural equation modeling, Brunner and colleagues (Brunner et al., 2010; Brunner, Keller, Hornung, Reichert, & Martin, 2009; Brunner, Lüdtke, & Trautwein, 2008) posited a new hierarchical and multifaceted model of academic self-concept, that is, the nested Marsh/Shavelson model (see Figure 2b). The nested Marsh/Shavelson model integrates the assumption of the hierarchical structure of self-concept by Shavelson et al. that one general academic self-concept factor is at the apex with the multifaceted nature of academic self-concept. In the model, subject-specific academic self-concept factors (e.g., subject-specific math academic self-concept or subject-specific verbal academic self-concept) are nested under a general academic self-concept factor representing the multidimensional nature of academic self-concept. The general academic self-concept factor exerts a direct influence on all indicators of general academic self-concept and its facets—that is, general as well as all subject-specific measures. Hence, the general academic self-concept factor is the most general factor in the structural model of academic self-concept, which is in line with its original conception in the Shavelson model (Shavelson et al., 1976). Subject-specific academic self-concept factors are residualized by the influence of general academic self-concept factor. Subject-specific academic self-concept factors are allowed to correlate. These correlations between the different subject-specific academic self-concepts indicate how students contrast their strengths and weaknesses in various subjects against each other, thereby the nested Marsh/Shavelson model offers a specific view of the structure, that is, students' academic self-concept profiles of subjective strengths and weaknesses in particular subjects. Further, the nested Marsh/Shavelson model accounts for the *shape* of students' profiles.

Brunner et al. (2010) found evidence for the nested Marsh/Shavelson model in a representative sample of eighth graders (N = 4,847) from Luxembourg. The nested Marsh/Shavelson model captures the structure of self-concepts in six core subjects best compared to the Marsh/Shavelson model (and a first-order factor model or general factor model). Further, evidence for a restricted version (using measures of general academic self-concept, math academic self-concept, and native language academic self-concept) was found in a cross-cultural sample of 15-year-old students in 26 countries, a restricted version of the nested Marsh/Shavelson model (i.e., general, math and native language self-concepts) accounted well for the structure of academic self-concepts (Brunner et al., 2009), and in representative data from 25,301 9th grade students from Luxembourg (Brunner et al., 2008). Moreover, Esnaola et al. (2018) found evidence for the restricted NMS model in Spanish secondary students comprising the grades 7 to 12 from two urban secondary schools and middle-class families. However, up to now, the nested Marsh/Shavelson model has never been investigated in elementary school students. To close this gap, the nested Marsh/Shavelson model is investigated in elementary school students in Study 2.

In the following, self-esteem, academic self-concepts, and social self-concepts are introduced and discussed with respect to the short-term and long-term consequences of self-concept development in childhood and adolescence for behavioral outcomes in education and in other areas of life.

2.1.1 Self-esteem

Self-esteem is defined as "the individual's positive or negative attitude toward the self as a totality" (Rosenberg et al., 1995, p. 141)[5]. Self-esteem becomes particularly salient in the vocabulary of young adolescence (Harter, 2012b, p. 79). Thus, adolescence is an important developmental period to investigate self-esteem. So far, much empirical work has indicated that self-esteem is positively related to happiness, general well-being, and mental health (cf., Baumeister et al., 2003; Marsh et al., 2017). For instance, several studies indicated that high self-esteem in adolescence is predictive for low depressive symptoms in adulthood (Ju & Lee, 2017; Kuzucu, Bontempo, Hofer, Stallings, & Piccinin, 2014; e.g., Orth, Robins, & Roberts, 2008; Rieger, Göllner, Trautwein, & Roberts,

[5] It should be noted that the terms self-esteem and self-concept are used in several studies interchangeably, (see, e.g., Riding and Rayner, 2001). In the present dissertation the term self-esteem is used for general self-concept, only.

2016). For example, using data from a 23-year longitudinal study (N =1,527) of German individuals from the LifE- study (Fend, Georg, Berger, Grob, & Lauterbach, 2002), Steiger et al. (2014) found that individuals at the age of 12 with high and/or increasing self-esteem in adolescence are found to be less prone to depression symptoms at age 16 as well as in adulthood (age 35). Thus, both self-esteem level at age 12 and mean-level change in self-esteem served as predictors for later depression symptoms. Further, Steiger et al. also investigated general academic self-concept and physical appearance and found the same pattern of results as with self-esteem. In a meta-analysis of longitudinal data from 77 studies on depression and 18 studies on anxiety comprising the age range of childhood to old age, Sowislo and Orth (2013) found that the effect of self-esteem on depression (β = 16) was significantly stronger than the effect of depression on self-esteem (β = .08); the effects between low self-esteem and anxiety were relatively balanced. Importantly, age did not function as a moderator so that this finding pertains also to childhood and adolescence. Further, individuals with depression in mid adolescence were at higher risk of educational underachievement in late adolescence and young adulthood (e.g., Fergusson & Woodward, 2002).

Moreover, individuals with a high self-esteem compared to those with low self-esteem show in part more favorable coping strategies with feedback after failure (Gibbons et al., 2002). Thus, self-esteem is significant in educational contexts because it plays an important role in students' coping with stress and strategies to deal with failures.

2.1.2 Academic self-concepts

Academic self-concepts refer to an individuals' self-evaluations of one's own abilities in school and different school subjects (e.g., Guay, Marsh, & Boivin, 2003). Academic self-concepts are important for educational attainment (e.g., Marsh et al., 2017). Given the multidimensionality of academic self-concept (see Section 2.1), the highest relations are found between achievement in a subject in a subject (e.g., math) and the matching subject specific self-concept (e.g., math self-concept) (see, e.g., Hansford & Hattie, 1982; Valentine, DuBois, & Cooper, 2004). Moreover, several studies indicated that academic self-concepts are longitudinally related to scholastic achievement (e.g., Marsh, Craven, & Martin, 2006). Furthermore, across different achievement levels academic self-concept is predictive for scholastic achievement (Susperreguy, Davis-Kean, Duckworth, & Chen, 2017). Moreover, prior academic achievement predicts subsequent self-concept across ages and gender. Thus, achievement is reciprocally related to scholastic achievement (Marsh & Craven, 2006; Marsh & Martin, 2011). Huang (2011) found evidence for the reciprocal effects model in a meta-analysis including 39 independent and longitudinal samples. Further, the majority of previous research indicated that the relation between achievement and academic self-concept is higher than those with self-esteem (Hansford & Hattie, 1982). This pertains also to the longitudinal relation (Marsh & O'Mara, 2008).

In addition, academic self-concepts predict academic behavior with long-term consequences, for example, coursework selection (e.g., Marsh & Yeung, 1997) and career choices (e.g., Guay, Larose, & Boivin, 2004), and approaches to learning (Burnett, Pillay, & Dart, 2003). Moreover, academic self-concepts mediated the relation between achievement and subject-specific interests (Schurtz, Pfost, Nagengast, & Artelt, 2014). Importantly, a positive development of academic self-concepts in adolescence seems to be also important for mental health in adulthood (Steiger et al., 2014).

2.1.3 Social self-concepts

Social self-concepts are self-perceptions with respect to social interactions with others and derive from the assessment of an individuals’ behavior within a given social context (Markus & Wurf, 1987). Social self-concepts refer to an individuals’ perceptions of social acceptance by others (i.e., social self-concept of acceptance) as well as to the evaluation of social competence in social interactions with others (e.g., social self-concept of assertion) (Berndt & Burgy, 1996). Like academic self-concept, previous research on social self-concept indicated that its structure is hierarchical and multidimensional, too. In line with the assumption that social selves are specific to a particular social contexts, across three age groups (pre to late adolescence) Byrne and Shavelson (1996) showed that a general social self-concept could be decomposed hierarchically into two major facets: Social self-concept as it relates to the family ("siblings" and "parents") and social self-concept as it relates to the school environment ("classmates" supplemented by "same sex peers" and "opposite sex peers" as well as "teachers"). Social self-concepts can also be differentiated with respect to the content (not context). For instance, Fend and Prester (1986) differentiated between social self-concept of acceptance and social self-concept of assertion. Social self-concept of assertion can be seen as a specific facet of social competence. Trautwein (2003) demonstrated that both facets are correlated but distinct from each other. Of note, the majority of studies focused on social self-concept of acceptance or social self-concept of competence as mixture of social acceptance, popularity, and the ability of 'how to make friends'. In the SDQs, social self-concept is assessed with a mixture of items including social acceptance, popularity, and the ability of 'how to make friends'. Recently, other self-concept researchers have also implemented aspects of perceived social competence in their measurement instrument (e.g., the revised version of the Self Perception Profile for adolescence; Harter, 2012). However, social self-concept of assertion as a separate facet of the social self-concept of competence has been widely neglected (Preckel, Niepel, Schneider, & Brunner, 2013; Young & Mroczek, 2003). Thus, in the present dissertation, both the social self-concept of acceptance and assertion were investigated in Study 1 (Chapter 3).

Social self-concepts play important roles in the social skill development of children and adolescents (see e.g., McElhaney, Antonishak, & Allen, 2008). Because social skills can be expected as general life skills (see Krapp, 2005), they are important in different areas of life. Further social self-concepts are related to social status, peer affiliations, and in the explanation of several internalizing (e.g., anxiety) and externalizing (e.g., aggression) symptoms (Hawley, 2003; Prinstein

& La Greca, 2002). For instance, McElhaney et al. (2008) showed, in a sample of $N = 164$ adolescents, that social self-concept of acceptance at age 13 predicted (even after controlling for peer status, i.e., sociometric popularity) at age 14 indicators of social functioning, that is, low aggression/hostility, desirability as a companion, and withdrawal. Trautwein, Köller, and Baumert (2004) studied German students in grade 7 and grade 10 and found that especially students with low self-concept of acceptance and assertion in grade 7 profited from engagement in acts of physical aggression with respect to self-concept of acceptance and assertion in grade 10.

Regarding the relation between social self-concept of acceptance and scholastic achievement, most previous research indicated that it is unrelated (e.g., Song & Hattie, 1985), weakly positive (e.g., Preckel, Zeidner, Goetz, & Schleyer, 2008), or weakly negative correlated (Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2006).

To conclude, all described self-concept facets have long-lasting effects in different areas of life. Self-esteem is just as important as academic self-concepts or social self-concepts for later psychological and behavioral adjustment. However, regarding the relation and prediction of academic achievement and academic career, the majority of studies indicate that academic self-concept is more predictive than self-esteem or social self-concepts.

2.2 Development of the self

In this chapter, important theoretical assumptions and models regarding the development of self-concept are described. Primarily, the theoretical work by Susan Harter regarding the development of the self is discussed because she offers a comprehensive theoretical view of self-concept development. Further, the theoretical assumptions of Shavelson et al. on stability and structural change and their refinements by Marsh and Ayotte (2003) in childhood and adolescence are outlined. Next, results from previous research are briefly described. Thereafter, descriptions of the different perceptions of stability and change in self-concept development are provided, and the findings from previous research relating these to the self-concept facets focused on in the present dissertation are summarized.

2.2.1 Development of the self-concept: Theoretical perspectives

Several self-concept researchers (Harter, 2012b; Marsh & Ayotte, 2003; Shavelson et al., 1976; Wigfield & Eccles, 2002) assume that self-concepts change with human development. These changes are in large part due to cognitive-developmental changes (i.e., individuals' cognitive capabilities) and to the interaction with the individuals' socialization experiences. These changes concern the attributes that define the self-concept, the organization, the stability, the accuracy, the determinants, and the ability to observe the self. Harter (2012b), referring to Piaget's *cognitive-developmental theory* (Piaget, Inhelder, Sinclair-De Zwart, & Pomerans, 1979), proposed a comprehensive description of the normative developmental liabilities and determinants of the self in childhood and adolescence. In particular, Harter (2012b) targets the development in several age ranges: very early childhood (ages 2 to 4), early to middle childhood (ages 5 to 7), middle to late childhood (ages 8 to 10), early adolescence (ages 11 to 13), middle adolescence (ages 14 to 16), and late adolescence (ages 17 to 19). Harter assumes that self-representations are unsystematically organized and unrealistically high until middle childhood. In early to middle childhood, the first kind of comparison processes influences self-perceptions (i.e., temporal comparisons, comparisons of attributes within a person across time). Further, an individual's grasp of self-concept changes from concrete descriptions of behavior to trait-like psychological constructs in middle childhood. Harter (2012) further assumes that domain-specific self-concept facets develop early; an integrated view (i.e., self-esteem) can first be verbalized at age 8 (but see, Harris, Donnellan, & Trzesniewski, 2017). In middle to late childhood, social comparisons (i.e., comparisons of attributes with other persons or groups) influence self-perceptions. In this period, children's cognitive capabilities allow the integration of information from social comparisons in their self-view. From middle to late childhood, children are able to integrate opposing but coexisting concepts into their self-concept (e.g., to be dumb and smart at the same time). During late child-hood, comparisons shift from those between different areas (e.g., students, friend) to comparisons between domains within the same area (e.g., in school: comparison of one's own ability in different school subjects), and to comparisons within a single domain (e.g., reading and writing in one's native language). Thus, with increasing age, dimensional[6] comparison processes (i.e., comparisons between domains) become more differentiated and more relevant for children's views of them-selves (Harter, 1999). Further, the accuracy of self-perceptions (i.e., the correlation between self-perceptions and external criteria) becomes higher. In early adolescence, individuals are capable of thinking abstractly ("I am intelligent").

Further, in early to middle adolescence, the term self-esteem becomes salient. Moreover, behavior and self-perceptions in different social contexts gain significance and, therefore, social skills are of great importance in this period. Social comparisons continue to influence self-perceptions. Early adolescence is a time in which school transitions occur (from elementary school to secondary school). These transitions imply changes in social contexts (i.e., new classrooms and classmates). Harter (2012) draws upon stage-environment fit theory (Eccles et al., 1993; Eccles & Roeser, 2009) in which it is assumed that educational classrooms and educational practices may show a mismatch with the developmental needs in early to middle adolescence. This mismatch likely leads to a decline in motivation, engagement, interest, and self-perceptions. Wigfield, Eccles, Mac Iver, Reuman, and Midgley (1991) found a decline during transition to middle school in self-esteem and academic self-concepts general math and verbal and social self-concept, but self-esteem and social self-concept (a mixture of social self-concept of acceptance and competence) recovered in

[6] It should be noted, that Harter do not use the term dimensional comparisons.

grade 7. In middle adolescence, individuals begin to relate one abstraction to another ("I am intelligent, and I am an airhead"). In late adolescence most individuals have transitioned through puberty, and cognitively, most adolescents have made the transition from concrete to formal operations which enables them to comprehend complex cognitive linkages (see e.g., Harter, 2012b). In this vein, Harter (2012) assumes that social comparisons now should have less influence on self-perceptions and the ideal standards of adolescents gain more impact on self-perceptions. Thus, there is an increasing impact of one's own ideal standards on perceptions.

Further, Harter (2012) assumes that self-perceptions also become increasing integrated with age, in that skills in different school subjects can be subsumed under the self-concept of being smart. She also assumes that self-perceptions become increasing differentiated with age, in that a global perspective of being smart generally develops into more differentiated self-representations in specific school subjects. This is in line with Shavelson et al.'s (1976) assumption that self-concept becomes more differentiated with age, which means that the hierarchy becomes weaker, that is, the relation between self-concept facets decrease with age (feature 7). Previous research indicated that self-concept becomes more differentiated with age for very young children, but suggests that this differentiation may plateau during late preadolescence when social comparison processes and cognitive abilities are adequately developed (see e.g., Marsh, 1989) (see the next section in detail). However, some parts of self-concept may become more integrated whereas other parts become more differentiated. "Theoretically, there is a need to integrate Harter's notion that self-concept becomes more integrated with age and the Shavelson et al. contention that self-concept becomes more differentiated with age" (Marsh, 1990b, p.109). Thus, Marsh and Ayotte (2003) occupy an intermediate position positing the differential distinctiveness hypothesis. Drawing on evidence from previous research regarding the prediction that self-concept becomes increasingly multidimensional with age (e.g., Marsh & Shavelson, 1985) was the assumption that correlations among all self-concept facets declined equally with increasing age, but previous research indicated that some correlations decline more than others (e.g., Marsh, 1989)

Marsh and Ayotte (2003) posit that:

> “With increasing age and cognitive development, there should be an increasing integration of closely linked areas of self-concept as well as increasing differentiation of disparate areas of self-concept. Whereas correlations among some self-concept factors should decline substantially, others should be expected to remain substantial or even increase with age” (pp. 689-690).

Marsh and Ayotte (2003) investigated grade 2 to grade 6 students (ages 6 to 12) and found evidence for their predictions. However, only reading self-concept was considered to assess native language self-concept and no other facets of native language self-concept like writing self-concept. Moreover, a first order factor model was used to investigate structural change and not the nested Marsh/Shavelson model. Thus, this is targeted in Study 2 (Chapter 4).

Further, Shavelson et al. assume that more general self-concept facets should be more stable than more specific ones (feature 4). However, evidence regarding this prediction is mixed; in most previous studies, the rank-order stabilities of general self-concepts are lower or about the same size as the rank-order stabilities of specific self-concepts (Marsh & Yeung, 1998; Shavelson & Bolus, 1982).

In the following different perspectives of stability and change in the development of self-concepts are described in detail.

2.2.1.1 Stability and change in the development of the self

Stability and change of self-concepts development can be considered and investigated under different perspectives (see Mortimer, Finch, & Kumka, 1982). Each highlight different aspects of the self-concept development and are conceptually and statistically distinct. Roberts, Wood, and Caspi (2010) arranged different perceptions of stability and change in a 2x2 table with sample level (named population level) vs. individual level and relative vs. absolute: rank-order consistency (relative, population), mean-level change (absolute, population), ipsative consistency (relative, individual), and individual differences in change (absolute, individual). Roberts et al. (2010) considered structural consistency as a prerequisite to investigate different aspects of development rather than an aspect of change and stability itself. However, structural consistency comprises two parts—measurement invariance and structural invariance. The present dissertation takes a differentiated perspective. Measurement invariance is expected to be a prerequisite for investigating different aspects of stability and change. Thus, it will be introduced in a separate section because the topic of measurement invariance concerns all studies in the present dissertation. Structural changes within self-concept formation can be expected as a test of structural invariance of the relations between self-concept facets across ages.

In the following, the different perspectives are outlined accompanied with the previous empirical research on the self-concept facets focused on the present dissertation. First, stability and change on sample level is introduced. Thereafter, an outline of structural change is given followed by a description of measurement invariance. Finally, individual-level stability and change are introduced.

2.2.1.2 Sample-level change

Change and stability on sample level can be divided into mean-level change and rank-order consistency *Mean-level change* reflects differences in the average value of a given construct in a specific group across time (e.g., McArdle & Nesselroade, 2014; Mortimer et al., 1982); it indicates how a group develops on average. Mean-level change can show different shapes, that is, linear or nonlinear (e.g., quadratic, cubic) (see, e.g., McArdle & Nesselroade, 2014). Rank-order stability reflects the degree to which the relative order of individuals on a given construct is maintained across time (e.g., Mortimer et al., 1982). It demonstrates individuals' change in rank change over time, that is, the stability of interindividual differences (e.g., Robins & Trzesniewski, 2005)

Existing results on the *mean-level change* in adolescents' self-esteem were inconclusive. Several studies indicated that mean levels were high in childhood, dropped during adolescence, and increased from adolescence throughout adulthood (Chung et al., 2014; Orth & Robins, 2014; Robins & Trzesniewski, 2005; Robins, Trzesniewski, Tracy, Gosling, & Potter, 2002; Soest, Wichstrøm, & Kvalem, 2016). Huang (2010) found, in his meta-analysis including 130 samples, that the average increase or decrease in adolescence depended on the self-esteem measure that was used. Fewer studies have investigated whether mean-level change showed a *linear or nonlinear shape*. A recent study by Birkeland, Melkevik, Holsen, and Wold (2012) indicated a nonlinear shape rather than a linear shape of mean-level change.

Regarding *mean-level change*, several studies indicated a decrease in general academic self-concept throughout adolescence (Archambault, Eccles, & Vida, 2010; Fraine, van Damme, & Onghena, 2007; Fredricks & Eccles, 2002; Rieger et al., 2017) that is nonlinear in shape (see e.g., Birkeland et al., 2012; Cole et al., 2001).

Studies investigating *mean-level change* of social self-concept of acceptance indicated an increase throughout adolescence (Cantin & Boivin, 2004; Cole et al., 2001; Young & Mroczek, 2003) or a decrease for a short time span (ages 12 to 14) (Bolognini, Plancherel, Bettschart, & Halfon, 1996). Previous research on the social self-concept of acceptance indicated that the shape of change might be nonlinear in nature (e.g., Cole et al., 2001). Only one study investigated mean-level change in the self-concept of assertion. Trautwein (2003) found an increase in the social self-concept of assertion throughout grade 7 in a sample of German students. To date, no study has yet investigated the shape of mean-level change in the social self-concept of assertion. Thus, it is an aim of the present dissertation to investigate the mean-level change and the shape of change throughout adolescence (Study 1; Chapter 3).

A meta-analysis of 50 studies indicated that *rank-order stability* in self-esteem in early childhood is low, and increases throughout adolescence and young adulthood (age 9 to 18 r = .39 -.61: Trzesniewski, Donnellan, & Robins, 2003). General academic self-concept increases in *rank-order stability* throughout elementary school (Guay et al., 2003). During adolescence, findings point to moderate to high rank-order stability (6-month stability grade 5 to grade 11 rs =.67-.85: Cole et al., 2001; 2-year stability between grade 7 and 9 r =.42: Gogol, Brunner, Preckel, Goetz, & Martin, 2016). Previous studies found high *rank-order stability* of social self-concept of acceptance throughout adolescence (grade 5 to grade 11 rs =.64-.85: Cole et al., 2001; 4-month stability in grade 7 rs =.50-63: Trautwein, 2003). Trautwein (2003) found that rank-order stability in social self-concept of assertion was moderate to high (4-month stability in grade 7; rs = .44-.51). However, research regarding social self-concept of assertion is scarce. Thus, it is an aim of the present dissertation to investigate the rank-order stability throughout adolescence (Study 1: Chapter 3).

2.2.1.3 Structural change

Structural change refers to the change of relations between self-concepts facets over time. Does the self-concept become increasingly differentiated or integrated with age? Different methods within a latent or manifest variable framework can be utilized to answer this question. If self-concept facets are modeled within a latent variable framework, structural change concerns the question if covariances and/or correlations (i.e., relations) between self-concept facets as latent factors are equal across time points[7] (Sass & Schmitt, 2013). The investigation of structural

[7] Note that if testing relations between factors, this concerns covariance structures. Mean structure can be incorporated (mean and covariance structures [MACS]) and latent means can be tested regarding invariance over time and/or groups. However, this test corresponds to the test of latent mean-level change which is expected as a separate development aspect (see Mortimer et al., 1982; Roberts et al., 2010).

changes in academic self-concept is a major target of the present dissertation. To date, previous research investigated structural change of the relations between self-concept facets as manifest scale score or latent factors within a first-order correlated factor model (e.g., Marsh & Ayotte, 2003).

Previous research, which investigated the entire self-concept with the SDQs, revealed that self-concept becomes multidimensional, that is, the average correlations decrease with age but only until preadolescence (e.g., Marsh, 1989) (see Section 2.4). With particular focus on the academic section of the self-concept, the correlation between math and reading self-concept (descriptive and evaluative component not separated) declined with age as well as when only the descriptive part was compared (Marsh, 1990e; see Marsh & Ayotte, 2003). In young children, math and reading self-concept are substantially related (e.g., grade 2, r= .65; Marsh & Ayotte, 2003) but weakly correlated in grades 3 to5 (r= . 26: Marsh & Ayotte, 2003; see also Arens et al., 2011).

Note that no study has investigated structural change within the academic section of self-concept using the nested Marsh/Shavelson model. As outlined in Section 2.1, this model fit to the theoretical assumption that individuals exert strength and weaknesses independent from an overall level of their academic ability. Therefore, it seems especially appropriate for tackling questions on the structural change in the sense of profile formation of academic self-concept because it allows to control for the general academic self-concept level in students' profiles, thus directly depicting students' academic self-concept profiles of subjective strengths and weaknesses in particular sub-jects.

To close this gap, we investigated structural change (i.e., profile formation) within the nested Marsh/Shavelson model and the first-order correlated factor model, too (Study 2; Chapter 4).

2.2.1.4 The topic of measurement invariance

Self-concept facets were most often assessed with a self-report questionnaire (Byrne, 1996). Questionnaires with multiple items are often used to study scores on latent factors (van de Schoot, Schmidt, Beuckelaer, Lek, & Zondervan-Zwijnenburg, 2015) as is also the case in the three studies in the present dissertation. Latent modeling of self-concept facets has the advantage that measurement error is controlled for compared to manifest modeling (e.g., calculating the sum of the items values to a score) (e.g., Sass & Schmitt, 2013). Moreover, typically, the structural models of the self-concept by Shavelson et al., the Marsh/Shavelson model, and the nested Marsh/Shavelson model are structural models that are translated into a factor model, that is, a latent variable framework to be tested. For example, does a structural model with one higher-order factor fit the data better than one with two higher-order factors? And is this the case in all groups and/or time-points? In the present dissertation related questions are targeted in Study 2 (Chapter 4) in which different structural models of academic self-concept are compared within and across grade-levels.

Note that specific forms of measurement invariance are prerequisite for hypotheses tests of latent factors and relations between those across groups and/or time points. Measurement invariance requires that the association between the items and the latent factors of individuals should not depend on group membership or time points (see e.g., Mellenbergh, 1989; Meredith, 1993). Measurement invariance comprises several forms. Configural measurement invariance means that the same structural model fits in all groups/time points equally well. If configural measurement invariance is confirmed, more restrictive forms can be tested. Metric measurement invariance (i.e., weak factorial invariance) concerns the equality of factor loadings and scalar measurement invariance (i.e., strong factorial invariance) as well as means. Strict measurement invariance, as the most restrictive form, requires comparable residual variances across time and/or groups. It should be noted that this form is seldom investigated (Meredith, 1993; Vandenberg & Lance, 2000). To investigate structural change, that is, relations (i.e., covariances and correlations) between self-concept facets over time or groups, metric measurement invariance is a prerequisite. This is the case in Study 2 (Chapter 4). Scalar measurement invariance is required if the aim is to test latent means, that is, mean-level changes. This is the case in Studies 1 and 2 of the present dissertation. Although the topic of measurement invariance is not new, not all studies have tested it as a prerequisite to investigate mean-level change or structural change (see e.g., Marsh & Ayotte, 2003). It should also be pointed out that not all previous studies used a latent variable framework to investigate mean-level change. In the three studies of the present dissertation, the self-concept facets were modeled in a latent variable framework. Further, the level of measurement invariance that is required was tested across groups or measurement waves prior to the test of structural change or mean-level change.

2.2.1.5 Individual-level change

Change on the individual level can be divided into *interindividual differences in intraindividual change*[8] (Nesselroade, 1992) and *ipsative stability. Interindividual differences in intraindividual change* concerns the variation regarding individual trajectories. Latent growth curve modeling offers a sophisticated approach to charting variability in individual trajectories (McArdle & Nesselroade, 2014). Although a construct may change on average on the sample level, some individuals may show an increase, others show a decrease, and some show no change at all. Individual-level change represents the magnitude of increase or decrease in a construct over time exhibited by an individual. The absence of mean-level change for a construct at a sample level may mask the fact that the construct increased for a sizable number of individuals but decreased

[8] Another classification in the literature is with respect to the methodological approaches to address separate families of questions: person-centered, variable-centered, and person-specific approaches. Variable-centered approaches assume that all individuals from a sample are and “averaged” parameters are estimate. Person-centered approaches consider the possibility that the sample might include multiple subpopulations characterized by different sets of parameters (see, e.g., Morin et.al., 2017; Howard & Hoffman, 2017). Given this classification, *interindividual differences in intraindividual change* are detected using variable-centered approaches.

for a comparable number of persons, thus canceling out mean change. Similarly, substantial rank-order stability can be present at the sample level, while considerable change is apparent at the individual level. *Interindividual differences in intraindividual change* has received considerable attention in the clinical psychology (Fryer & Elliot, 2007) and has received a less attention in the self-concept research (Robins, Fraley, Roberts, & Trzesniewski, 2001). Latent growth curve models (e.g., Nesselroade, 1992) and the reliable change index (RCI; Jacobson & Truax, 1991) can be used to examine individual-level change[9].

Only a few studies investigated interindividual differences in intraindividual change in multiple self-concept facets in one study. They found significant variability in self-esteem, academic self-concept, and social self-concept of acceptance (see, e.g., Young & Mroczek, 2003). However, regarding social self-concept of assertion, evidence is totally lacking in previous research. Thus, in Study 1 (Chapter 3), social self-concept of assertion from late childhood to late adolescence is investigated.

Despite *interindividual differences in intraindividual change,* ipsative stability describes intraindividual ordering of different self-concept facets over time; thus, the change in an individual's self-concept profile across time. Ipsative stability comprises three aspects: elevation (the level of profile scores), shape (the pattern of profile scores), and scatter (the variability of profile scores) (Fryer & Elliot, 2007).

Often ipsative stability is tested using Cronbach and Gleser's (1953) D^2 statistic or the q-correlation (i.e., the product-moment correlation of individual profiles) (see Trucco, Wright, & Colder, 2014). However, more recent statistical approaches are mixture modeling methods like latent transition analysis (see Wright & Hallquist, 2014). Note, that ipsative stability is not explicitly targeted in

[9] The RCI allows to categorized individuals whether showing a significant increase, a significant decrease, or a significant change between time points. It is also possible to aggregate these data to make summary statements about reliable change within the sample of individuals as a whole (see Jacobson and Truax, 1991).

the present dissertation but its description is included for the sake of completeness. Further note that although we investigate the development of the *profile* of academic self-concept using the nested Marsh/Shavelson model (an incomplete bifactor model) in Study 2, the chosen approach still makes statements on a sample level concerning the structure of self-concepts and structural changes[10].

After the introduction of theoretical assumptions of the developmental of self-concept aspects and different perceptions of stability and change in self-concept development accompanied by the status of previous research in the following theoretical assumptions and findings from previous research regarding whether changes in one self-concept facet are causally longitudinal related to changes in another self-concept facet are outlined.

2.3 Causal flow between self-concept facets: Theoretical perspectives

It is a matter of debate in self-concept research how and if self-concept facets are causally longitudinal related (Hattie, 2004; Marsh & Yeung, 1998; Trautwein, 2003; Wagner & Valtin, 2004). Regarding the hierarchically structured Shavelson et al. model of self-concept, the causal flow between global self-concept (i.e., self-esteem) and between specific self-concept facets (e.g., academic and social self-concepts) can be bottom-up (former specific self-concept facets cause later general self-concept facets), top-down (former general self-concepts cause later specific self-concept facets), or reciprocal (i.e., bottom-up and top-down at the same time) (see Section 2.1). The longitudinal relations between specific self-concept facets are called transdimensional (Trautwein, 2003).

[10] Note that latent factor scores for each person estimated in factor analysis can be saved and an individual's profile of academic self-concept can be descriptively inspected. Further, latent factor scores derived from a bifactor model can be used in person-centered approaches like latent transition analysis to identify groups of individuals with different profiles (in level and shape) (see Morin et al., 2017).

Several cross-sectional studies with the SDQs (e.g., Arens & Morin, 2016; Byrne & Gavin, 1996; Byrne & Shavelson, 1996; Marsh & Ayotte, 2003) showed that self-esteem and self-concept facets are related. Across adolescence, the concurrent correlations between self-esteem and general academic self-concept or social self-concept of acceptance are positive (see e.g., Byrne & Gavin, 1996; Byrne & Shavelson, 1996; Marsh et al., 1988; Tetzner, Becker, & Maaz, 2016; Wagner & Valtin, 2004). Academic and social self-concepts showed small to moderate positive correlations (e.g., Byrne, 1996). Social self-concept of acceptance and social self-concept of assertion are highly positive correlated in grade 7 and in grade 10 (see Trautwein, 2003). However, these findings do not allow conclusions to be drawn on whether changes in one self-concept are causally related to changes in another self-concept.

In Section 2.3.1 theoretical assumptions and findings from previous research regarding the causal flow between self-esteem and specific self-concept facets are described and in Section 2.3.2 the causal interplay between specific self-concept facets are focused. Thereby, only findings from previous studies are presented that investigated the focused self-concept facets in the present dissertation (i.e., self-esteem, academic self-concepts and social self-concepts).

2.3.1 Causal flow between general and specific self-concept facets

Shavelson et al. (1976) assume that individuals' self-concepts are formed through feedback from significant others and social interactions with the environment. Therefore, the causal flow in the self-concept hierarchy is mainly from the bottom to the top so that, for example, success/failure in school leads to changes in academic self-concept that in turn cause changes in self-esteem. Similar predictions were made by Harter (1999, 2012b), who assumes that self-esteem results from one's accomplishments in important domains (e.g., academic). Further, feedback obtained from significant others is supposed to be an important determinant of self-esteem. However, Harter

(1999) and Shavelson et al. (1976) both acknowledged that generally it is, to a small degree, possible that self-esteem can influence domain-specific self-concepts. Therefore, Marsh and Yeung (1998) interpreted the assumption implied in the Shavelson model to be *reciprocal* in nature. Reciprocal relations mean that bottom-up and top-down effects can be found at the same time.

Contrary to Harter (1999) and Shavelson et al. (1976), Brown (1993; see also Dutton & Brown, 1997) assumed that the causal flow mainly goes from self-esteem to specific self-concepts, that is, *top-down*. Brown (1993) argued that “through transfer of affect processes and halo effects, positive feelings toward the self in general color people’s evaluations of their specific attributes” (p. 31). Brown assumes that self-esteem is learned in early years from unconditional love by parents. Only the case of extreme failure or success in an individually important domain of life may lead to a change in self-esteem. Brown (1993) found evidence for his assumptions by investigating mainly college students using a quasi-experimental design. However, only longitudinal studies allow researchers to draw conclusions about the directionality of relations between self-esteem and specific self-concepts (Marsh & Yeung, 1998). Some self-concept researchers (see Marsh & Yeung, 1998) assume that there is no exchange between general and specific self-concepts or only to a very small degree. Marsh and colleagues (Marsh, 1993b; Marsh & Yeung, 1998) assume that specific self-concepts have little predictive value for the development of self-esteem.

To summarize, the different assumptions are partly opposite of each other and partly contradicting. Thus, the pattern of bottom-up or top-down effects or reciprocal effects remains unclear. In this regard, Marsh and Yeung (1998) stated that “theoretical accounts of the direction of causal flow, however, have been ambiguous, have not adequately operationalized terms, and have not provided any defensible methodological approaches” (p. 511). Suls (1993) concluded, "both Brown and Marsh, who cite strong support for their viewpoints, cannot be right; or, at minimum, a new integrative theory is needed to reconcile the two approaches" (p. X).

Only few studies investigated the direction of longitudinal interplay between different self-concept facets using a longitudinal design[11]. For example, Trautwein (2003; Study 2) studied the longitudinal relations between self-esteem, academic self-concept in math and German, and the social self-concepts of acceptance and assertion across three time points in a sample of German 7th graders. He found some evidence for bottom-up effects with regard to math academic self-concept and language academic self-concept on self-esteem (range β = .07-.16). Further, he found weak evidence for top-down effects of self-esteem on specific self-concepts. Other studies found reciprocal relations between academic self-concept and self-esteem in middle adolescence (see e.g., Skaalvik & Hagtvet, 1990).

In sum, studies are inconclusive regarding the predominance and the pattern of causal longitudinal (bottom-up, top-down, reciprocal) relations between self-esteem, academic self-concepts, and the two social self-concepts. Overall, effect sizes of significant relations between self-concepts were small in most of the studies (but see Wagner & Valtin, 2004). Most studies covered only timeframes of a few month or years during adolescence (one year, 3 measurement waves: Trautwein, 2003; two years, 2 measurement waves: Wagner & Valentin, 2004).

2.3.2 Causal flow between specific self-concept facets

Effects between specific self-concepts at the same level or different levels of specificity are called *transdimensional* (Trautwein, 2003). The causal flow between specific self-concept facets can be unidirectional or reciprocal. Transdimensional effects are frequently investigated with respect to the influence of dimensional comparison processes in the academic section of the self-concept (see the following Section). Transdimensional effects between academic and social self-concepts and within social self-concepts were scarcely discussed and investigated (Trautwein, 2003). One of the

[11] Although Marsh and Yeung (1998) acknowledge that the term causal is only to be used in experimental designs.

few studies investigating transdimensional effects was conducted by Preckel, Niepel, Schneider, and Brunner (2013)[12]. They investigated German 5th and 8th graders and questioned if a positive development of academic self-concept is compatible with a positive development of social self-concept. They found that earlier social self-concept of acceptance negatively predicted changes in general academic self-concept over time while earlier social self-concept of assertion positively predicted changes in academic self-concept. However, effect sizes were small. Findings also indicated that social self-concepts were not longitudinally related to each other.

To conclude, transdimensional effects between general academic self-concept and the two social self-concepts or within social self-concepts were rarely investigated (Trautwein, 2003). Moreover, it is difficult to compare the existing studies because study designs varied in timespans and the self-concepts under consideration. Moreover, no study covered the whole timeframe of adolescence. Thus, conclusions on the causal flow between specific self-concept facets throughout adolescence remain elusive. To close this research gap, in Study 1 (Chapter 3) we investigated the causal flow between self-esteem, general academic self-concept, and social self-concept of acceptance and assertion from grade 5 to grade 11, covering late childhood, and early, middle, and late adolescence. Furthermore, we aimed to replicate and extend the previous study by Preckel et al. (2013) with an extended sample including more time-points and including self-esteem in the analysis.

In the following, the focus shifts on the academic section of the self-concept. Comparison processes as determinants for subject-specific academic self-concept development and academic self-concept formation are described and research findings are presented.

[12] Preckel et al. (2013) used a partly overlapping and a partly extended sample of the sample included in the present study. Our sample differed in two ways from that of Preckel et al. (2013). First, two cohorts of students from one school in the sample were excluded because this school changed the length of schooling from nine years of secondary education (after grade 4) to eight. Second, we included two additional measurement points in grade 5, and because the sample de-rive from an ongoing project, one assessment wave in grade 8 and data from grade 11 were added.

2.4 Determinants of academic self-concept: Comparison processes

As major determinants of self-perceptions comparison processes within different frame of references are expected as important determinants for the development of self-concepts across different theories (cf., Möller & Marsh, 2013) (see also the theoretical perspective of self-concept development by Harter [2012] in Section 2.4). If an individual evaluates his/her ability/achievement, this does not take place in a vacuum. Accomplishments from different sources are usually compared against frame of references (Marsh, 1990c, 1993a). Motivation for self-evaluations are manifold; they can support self-verification, self-enhancement or self-maintenance (cf., Möller & Marsh, 2013).

Three types of comparison processes can be distinguished, namely social, temporal, and dimensional. First, one of the most targeted comparisons of self-perceptions are social comparisons. Social comparisons refer to comparison with other individuals or groups, they are interindividual and within an external frame of reference. Although they develop very early in life, they become integrated in individuals' self-view in middle childhood (see Section 2.2.1). Second, temporal comparisons refer to comparisons of attributes within an individuum across time, they were intraindividual and within an internal frame of reference. They develop very early in life before social self-concept is used in early childhood (see Section 2.2.1). Both kinds of comparison processes are well established and elaborated in psychology: social comparisons theory (e.g., Festinger, 1954) and its refinements (e.g., Corcoran, Crusius, & Mussweiler, 2011) as well as temporal comparison theory (Albert, 1977) (see Möller & Marsh, 2013).

Recently, educational psychology has focused on a third type of comparison processes, that is, dimensional comparisons. It is expected that individuals conduct dimensional comparisons from late childhood onwards (see Section 2.2.1). Dimensional comparisons are intraindividual comparisons within an internal frame of reference in which an individual compares different attributes of themselves with other attributes of themselves. For example, an individual compares his/her achievement in math with his/her achievement in the native language subject. This type of comparisons are explicitly focused[13] in a prominent model for the development of subject-specific academic self-concepts and the academic self-concept formation, the so-called I/E model (Marsh, 1986). Hence, in the following, the I/E model is described in detail. In the meantime, the assumptions of the I/E model was generalized in the dimensional comparison theory (Marsh et al., 2014; Marsh et al., 2015; Möller & Marsh, 2013) that is also outlined in the following section.

2.4.1 Dimensional comparison processes: The I/E model

Dimensional comparison processes are assumed to be the reason why academic self-concepts are multidimensional with separate, domain-specific self-concepts for various achievement domains such as specific school subjects (e.g., Brunner et al., 2010). Especially, math self-concepts and verbal (native language) self-concepts were found to be nearly uncorrelated or only weakly correlated (e.g., Marsh, 1990b, 1990c) (see Section 2.1). This finding strongly contrasts the ubiquitous finding that those students with strong (weak) achievement in math typically also demonstrate strong (weak) verbal achievement – a correlational pattern among achievement measures that suggests substantial correlations between corresponding self-concepts (Marsh, 1986, 1990b, 1990c).

To explain these paradoxical pattern the I/E model (Marsh, 1986) was developed. The I/E model was tested with cross-sectional, longitudinal, experimental, and introspective studies. Evidence for the predictions within the I/E model came from a meta-analysis (Möller & Köller, 2004; Möller, Pohlmann, Köller, & Marsh, 2009). The meta-analysis comprised 69 data sets (N =125,308).

[13] The term dimensional was introduced by (Möller and Köller, 2001)

The I/E model assumes that two different comparison processes are the major determinants for the development of subject-specific self-concepts of students, namely social comparisons within an external frame of reference and dimensional comparisons within an internal frame of reference. Within an external frame, students compare their achievement/ability with their peers (e.g., classmates) and related information such as grade distribution (i.e., social comparisons) (e.g., Marsh et al., 2014; Möller et al., 2009). If students' subject-specific achievements are higher than the class-average achievement, the subject-specific self-concept will also be higher (Möller et al., 2009, p. 1131). Further, a student with the same ability (as measured by standardized tests) has lower academic self-concepts when they attend higher ability classes or schools than when they attend lower ability classes or schools, a well replicated effect known as the big fish little pond effect (BFLPE; Marsh, 1987, 2005; for recent reviews, see Marsh & Seaton, 2015). Within an internal frame of reference, students compare their achievement/ability in one subject with their (perceived) achievement/ability in another subject (i.e., dimensional comparisons) (e.g., Marsh, 1986; Marsh et al., 2014; Möller et al., 2009). It is assumed that students compare/contrast their achievement in math with their achievement in the native language. Thus, a negative correlation between math and native language self-concept can be expected as a result of dimensional comparisons. As achievements between subjects are usually substantially positively correlated, a positive correlation can be expected as a result of social comparisons. The joint operation of these two counterbalancing comparison processes explains the small or near-zero correlations between math and native language self-concept.

Thus, the I/E model assumes that since math and reading ability/achievements are compared with each other, it is the difference between math and verbal skills that contributes to a higher self-concept in one area or the other.

Findings regarding the I/E model were generalized in the *dimensional comparison theory* (Marsh et al., 2015; Möller & Marsh, 2013). Within the dimensional comparison theory, the I/E model was extended by applying it to more than two domains. It is assumed that the perceived similarity between subjects moderated the effect of dimensional comparison which could be confirmed in an experimental study (Helm, Mueller-Kalthoff, Nagy, & Möller, 2016). Subjects can be aligned along a continuum that is derived from the different correlations of subject-specific self-concepts in the Marsh/Shavelson model (see Figure 2a). Contrast effects are more likely for far subjects than for near subjects (e.g., math and native language).

Assimilation effects are more likely the nearer the subjects along the continuum are (e.g., native language and foreign language). Regarding the extension to more than math self-concept and verbal self-concept, several studies investigating further subjects concepts (Arens, Möller, & Watermann, 2016; Jansen, Lüdtke, & Schroeders, 2016; Jansen, Schroeders, Lüdtke, & Marsh, 2015; Marsh et al., 2015; Rost, Sparfeldt, Dickhäuser, & Schilling, 2005; Schilling, Sparfeldt, & Rost, 2004) have shown, for instance, assimilation effects between native language and foreign language (Jansen et al., 2015). In the meantime, dimensional comparison theory was extended to a broad range of other domain-specific constructs like motivation, learning behavior, or personality characteristics (cf., Möller, 2016). For instance, Möller and Savyon (2003) investigated the effect of dimensional comparisons regarding the academic and non-academic self-concept (i.e., honesty) facets. In N= 70 high school students (aged 12 to 19) they showed that low self-concept students rated their honesty slightly more positive than students with a high academic self-concept. Note that the longitudinal relations between academic and social self-concepts can therefore also be discussed as being a result of dimensional comparison processes (see Subsection 2.3.2).

Despite of these current development regarding dimensional comparison theory it is the question of which factors moderate the effects of dimensional comparisons within the original I/E model that focuses math and native language self-concepts. These are reviewed in the following section.

2.4.1.1 The I/E model: Moderators

Several studies investigated moderator variables concerning the effect of dimensional comparisons between math and native language self-concepts. For instance, Möller et al.'s (2009) meta-analysis of the I/E model indicated that the kind of self-belief measure (self-academic self-concept vs. efficacy), the kind of achievement measure (grades vs. standardized achievement tests), the year of publication, the sample size, and the years in school (i.e., age) all act as a moderator whereas gender and country did not. Note, however, that only three samples of elementary school students were included in this meta-analysis. As most research to date has focused on secondary school students, evidence for elementary school students is comparably scarce. To close this research gap, the I/E model is investigated in elementary school students in grades 3 and 4 in Study 2 (Chapter 4).

Further, Steinmayr and Spinath (2015) found that intelligence moderated the effects of dimensional comparisons. Dimensional comparison effects were generally more pronounced for more intelligent students than for those of lower intelligence in a sample of 11th and 12th grade German students.

Besides intelligence as a moderator, the specificity of ability beliefs, that is, of whether an individual perceived that their math and verbal abilities are correlated or not has found to be a moderator of dimensional comparison processes. (Möller, Pohlmann, Streblow, & Kaufmann, 2002) investigated German secondary students (7th and 9th graders) and found that the effects of dimensional comparisons were stronger for students who perceived their math and verbal ability to be uncorrelated.

Moreover, the size of the grade difference was found to be a moderator of the effect of dimensional comparisons as well. Rost et al. (2005; see also Schilling et al., 2004) showed, in a sample of

German secondary school students (7th and 8th graders), that the higher the size of the grade difference, the stronger the effects of dimensional comparisons were.

To summarize, several moderator variables of dimensional comparisons have been investigated but predominately in secondary school students whereas this research area has been widely neglected in elementary school students. To close this research gap, we investigated if grade difference acts as a moderator in elementary school students in Study 2 (Chapter 4). The lower the grade difference, the lower the correlation between math and native language self-concept should be.

2.4.1.2 The I/E model: Statistical approaches

Aside from the topic of which factors are moderators in the I/E model, there are different statistical approaches to investigate the I/E model when the data are cross-sectional and nonexperimental in nature. Up to now, most evidence for the I/E model stem from cross-sectional data and nonexperimental research designs (see, e.g., Möller et al., 2009). These studies mainly used path analytic approaches. The majority of studies applied a path-analytic approach, in which math self-concept is regressed on math and verbal achievement and verbal self-concept is regressed on math and verbal achievement (e.g., Marsh, 1986). Social comparison processes imply positive paths from math achievement to math self-concept and from verbal achievement to verbal self-concept reflecting that a student's higher achievement in one subject relative to his or her peers is related to a higher level of self-concept in the same subject. Dimensional comparison processes, on the other hand, imply negative paths from math achievement to native language self-concept and from native language achievement to math self-concept reflecting contrast effects of an improvement in achievement in one subject on the self-concept in the other subject

However, the Statistical Inference of the American Psychological Association states that "if the assumptions and strength of a simpler method are reasonable for your data and research problem, use it” (Wilkinson, L. & Task Force on Statistical Inference, American Psychological Association, Science Directorate, 1999, p. 598). With this in mind, there is an alternative approach to investigate the I/E model with cross-sectional and nonexperimental data proposed by Rost et al. (2004; 2005). They used a quasi-experimental approach to investigate the I/E model. Because this designation is somewhat misleading, we rename this approach the correlation comparison approach (CCA). This seldom-used approach actually offers different statistically favorable advantages, for example, no susceptibility to multicollinearity. Moreover, a key assumption of the I/E model is directly addressed. That is, since math and reading ability/achievements are compared with each other, it is the difference between math and verbal skills that contributes to a higher self-concept in one area or the other. To close this research gap, we used the CCA to investigate the I/E model in Study 3 (Chapter 5).

2.5 The present study

The research gaps outlined in Chapters 1 and 2 on the structure and the development of self-concepts and their determinants in childhood and adolescence were the subject of investigation in three studies.

Study 1 (Chapter 3) focused on the different aspects of development: on the sample level (mean-level, rank order) and individual level (interindividual differences in intraindividual change) of self-esteem, general academic self-concept, social self-concept of acceptance, and assertion from late childhood to late adolescence. Further, we investigated the pattern of the causal flow between these self-concept facets.

In Study 2 (Chapter 4) and Study 3 (Chapter 5), the focus shifts to the academic section of the self-concept.

In Study 2 (Chapter 4), the recently proposed structural model of academic self-concept (i.e., the nested Marsh/Shavelson model of academic self-concept) is tested in elementary school students. Further, mean-level change and structural change (i.e., differentiation and integration) of academic self-concept formation in elementary school students are investigated. As major determinants of academic self-concept formation, the influence of dimensional comparisons and social comparisons as posited in the I/E model is tested. Moreover, we examine grade difference as a moderator of dimensional comparisons.

In Study 3 (Chapter 5), we tested whether dimensional comparisons depend on the grade difference (i.e., a key assumption of the I/E model), and the I/E model is investigated using a seldom-used statistical method. Going further, our samples comprised elementary school students and secondary students (grades 3 to 8). Therefore, not only is the generalizability of I/E model findings over the used method tested in this study but also the different age ranges and self-concept scales.

Table 1 provides a detailed overview of the self-concept facets, measurement instruments, age range, data source, study design, and the considered perceptions of change and stability in self-concept development that were targeted in the each of the present studies.

Table 1. Overview of the three studies

	Study 1	***Study 2***	***Study 3***
	Developmental dynamics of adolescent students' self-esteem, academic self-concept and social self-concepts: Results from a six-year longitudinal study	***Profile formation of academic self-concept in elementary school students in grades 1 to 4***	***Effects of achievement differences for internal/external frame of reference model investigations: A test of robustness of findings over diverse student samples.***
Self-concept facets	Self-esteem, general academic self-concept, social self-concepts of acceptance and assertion	General academic self-concept, subject-specific academic self-concept in math and the native language (reading, writing)	Subject-specific academic self-concepts in math and the native language
Assessment	Questionnaire: selected items SDQ-II, Items Rosenberg scale (1965), Fend & Prester (1986), grades	Questionnaire: FEESS-K 1-2 & 3-4 (Baudson & Preckel, 2015); grades - report cards	Questionnaire: selected items SDQ-II
Sample(s)	German secondary school students (Grades 5 [3 times], 6, 8, 11; ages: 8-20)	German elementary school students (Grades 1-4; ages 6-11)	German elementary school students and secondary school students (Grades 3-8; ages 8-16)
	N = 1,163	Grade 1: N = 427; Grade 2: N; = 718; Grade 3: N = 1,288; Grade 4: N = 1,346	Grade 3: N = 326; Grade 4: N = 808; Grade 5; N = 878/409; Grade 6: N = 836/461; Grade 7 N = 409; Grade 8 N = 841
Study design	longitudinal	cross-sectional	Elementary school students: cross-sectional; secondary school students: longitudinal but cross-sectionally investigated
Data source	AVG project	THINK project	THINK project[a]: elementary school students; AVG[b] /PULSS project: secondary school students
Perceptions of stability and change			
Mean-level change	X	X	X
Rank-order stability	X		
Individual-level change	X		
Structural change		X	

Note. [a]Sample size differed between Studies 2 and 3 because general academic self-concept was not included in the analysis in Study 3. For several students only general academic self-concept was assessed. Thus, due to handling of missing data using pairwise deletion in Study 2 the sample is larger in this study. [b]Students from Bavaria were excluded to obtain samples that do not suffer from an overlap regarding the same federal states of Germany.

3 Study 1

Developmental dynamics of adolescent students' self-esteem, academic self-concept and social self-concepts: Results from a six-year longitudinal study[14]

[14] Schmidt, I., Brunner, & Preckel, F. Manuscript prepared for submission

3.1 Abstract

In adolescence, the maintenance of a positive self-concept seems to be particularly important but also challenging. The present longitudinal study focused on the development of multiple self-concept facets during adolescence (i.e., self-esteem, academic self-concept, social self-concepts of assertion and of acceptance). Over a six-year period, change on the sample-level (i.e., rank order stability, mean-level change), change on the individual-level (interindividual differences in intra-individual change) as well as longitudinal relations between the self-concept facets were investigated. The sample comprised $N = 1{,}163$ German students (46.7% female) attending the highest track of the German secondary school system who were followed from grade 5 to grade 11 in six waves of measurement. Data analytic methods involved second-order latent growth curve models and structural equation modeling. All self-concept facets became more stable in rank-order) over time and were comparably stable. But, shapes of mean-level change differed between facets. Students' individual developmental trajectories showed significant variability for all self-concept facets. Longitudinal relations between self-concept facets were small, indicating effects from the earlier social self-concept of assertion and academic self-concept on later self-esteem, effects from earlier self-esteem on later academic self-concept, and reciprocal relations between academic self-concept and self-esteem. Practical implications regarding self-concept interventions in adolescence were discussed.

Keywords: self-esteem; academic self-concept; social self-concept of assertion; social self-concept of acceptance; adolescents; stability; interindividual differences in intraindividual change

3.2 Theoretical background

In adolescence, positive beliefs about oneself (i.e., self-concepts) are important for a positive development in general and in different areas of life such as academics or social relations (Baumeister, 1999; Baumeister et al., 2003; Harter, 2012b; Marsh, Craven et al., 2006). Adolescence is a developmental period where it seems particularly challenging to maintain or develop positive self-concepts because of the many psychological, physical, and contextual changes and transitions that occur during that time (e.g., Harter, 2012b). Major developmental challenges are faced in the physical, academic, or social domain (see, e.g., Havighurst, 1979). Mastering these challenges has meaningful implications for later adjustment in adulthood (see, e.g., Orth, Robins, & Widaman, 2012; Steiger et al., 2014)

When investigating the development of self-concepts in adolescence, the question arises if all or most individuals show a comparable development or if there are interindividual differences in development. The answer to this question bears important information for the design of interventions (e.g., general vs. group-specific interventions). Furthermore, many studies documented that self-concepts are multidimensional and hierarchically structured, comprising general facets like self-esteem as well as domain-specific facets like academic or social self-concepts (e.g., Baumeister, 1999; Harter, 2012b; Marsh, Craven et al., 2006; Shavelson et al., 1976). Accordingly, another question arises, namely if different self-concept facets display a similar or different development. Again, the answer to this question has practical implications, informing practitioners about whether interventions should target multiple self-concepts or specific self-concept facets. However, up to now, only few studies investigated interindividual differences in intraindividual change while considering multiple self-concept facets in one sample using longitudinal data. Often, studies that investigated the development of multiple self-concepts did not focus on both the sample level and the individual level. Moreover, whereas studies frequently covered only a few month or years during adolescence when investigating stability and change in multiple self-concept facets (3 years; e.g., Bolognini et al., 1996; Granleese & Joseph, 1994) we were able to investigate this development over a six-year period (but see Kuzucu et al., 2014; Soest et al., 2016).

In the present study, we longitudinally investigated adolescents' developmental of several facets of self-concept, namely self-esteem, academic self-concept, and the social self-concepts of acceptance and assertion. Taking into account the developmental challenges in adolescence (see, e.g., Havighurst, 1979), it seems particularly important to focus on these self-concept facets. Self-esteem represents an individual's global evaluation of his or her overall worth as a person (Harter, 2012b; Rosenberg, 1965). Self-esteem is positively related to general well-being and mental health (see e.g., Marsh et al., 2017) and can be perceived as an indicator of psychological functioning (see e.g., Kernis, 1993). For example, individuals with higher self-esteem in adolescence are found to be less prone to depression symptoms in adulthood (Steiger et al., 2014).

Academic self-concepts refer to an individuals' self-evaluations regarding one's own abilities in school and in different school subjects (e.g., Guay et al., 2003). Academic self-concepts drive scholastic achievement (e.g., Marsh, Craven et al., 2006) and predict academic behavior with long-term consequences on, for example, coursework selection (e.g., Marsh & Yeung, 1997) and career choices (e.g., Guay et al., 2004).

Finally, social self-concepts refer to an individuals' perceptions of social acceptance by others (i.e., social self-concept of acceptance) as well as to the evaluation of social competence regarding social interactions with others (i.e., social self-concept of assertion) (Berndt & Burgy, 1996). Social self-concepts play important roles in youths' social skill development, social status, peer affiliations, and in explaining several internalizing (e.g., anxiety) and externalizing (e.g., aggression) symptoms (Hawley, 2003; Prinstein & La Greca, 2002).

Despite of the development of multiple self-concept facets, a major concern is how these self-concept facets are related across time. That is, does (change in) one self-concept facet causally predict (change in) another self-concept facet? Again, this knowledge is helpful to determine whether interventions designed to enhance self-concept facets produce side-effects on other self-concept facets. There is an ongoing debate on how self-concepts are causally related over time but there are only few longitudinal studies investigating this question and their findings are inconsistent (see e.g., Marsh & Yeung, 1998; Trautwein, 2003). Further, unlike most studies that covered only a few month or years during adolescence (one year 3 time points: Trautwein, 2003; two years 2 time pointsWagner & Valtin, 2004) we have the opportunity to investigate adolescent development over a six-year period.

To conclude, in the present paper we investigate the development of adolescents' self-concepts in vital domains (i.e., self-esteem, academic self-concept, and the social self-concepts of acceptance and assertion) over a six-year period. To get a more nuanced picture of individual differences in development and of causal relations between the various self-concept facets, we investigate development on a sample level and on an individual level. Furthermore, we examine the causal flow of the four self-concept facets.

In the following, we describe theoretical assumptions and empirical findings from previous research on the development and on the causal interplay of the four self-concept facets during (early, middle, and late) adolescence.

3.2.1 Self-concept in adolescence: Structure and development

One of the most influential structural models of the self was proposed by Shavelson et al. (Shavelson et al., 1976). This model posits a multidimensional and hierarchical structure of the self, which is composed of general self-concept (i.e., self-esteem) at the apex, and academic and three nonacademic self-concept facets (i.e., social, emotional, and physical self-concept) at a lower level of hierarchy. Of note, academic self-concepts and social self-concepts are hierarchical and multidimensional in their own right. Academic self-concepts can be structured into a general academic self-concept at the most general level and lower-order subject-specific self-concepts such as academic self-concept in math or physics (Brunner et al., 2010).

Regarding social self-concepts, Byrne and Shavelson (1996) investigated adolescents and found that a global social self-concept could be decomposed hierarchically into two major facets: Social self-concept as it relates to the family ("siblings" and "parents") and social self-concept as it relates to the school environment ("classmates" supplemented by "same sex peers" and "opposite sex peers" as well as "teachers"). With respect to the content (not context), Fend and Prester (1986) differentiated between social self-concept of acceptance and social self-concept of assertion. This is in line with the definition provided by Berndt and Burgy (1996) that social self-concepts refer to an individual's perceptions of social acceptance by others (i.e., social self-concept of acceptance) as well as to the evaluation of social competence regarding social interactions with others (i.e., social self-concept of assertion). In this respect, most studies focused on social self-concept of acceptance while social self-concept of assertion as a separate facet of social self-concept has been neglected (Preckel, Niepel, Schneider, & Brunner, 2013; Young & Mroczek, 2003). However, for positive development it is important to feel socially accepted *and* socially competent (Krapp, 2005). Competence in social interactions is perceived as a major social challenge faced by youths but nonetheless is needed for them to build and maintain positive peer relations in adolescence (Harter, 1999, 2012b). In the present study, we therefore investigated both adolescents' social self-concept of acceptance and their social self-concept of assertion.

In the following, we briefly introduce different conceptualizations of development that highlight the different aspects of self-concept development. It is important to understand the meaning of these different perspectives on change or stability for the investigation of the development of self-concept. Afterwards, we first describe theoretical assumptions on self-concept development in adolescence, followed by an overview of the empirical findings.

The *rank-order stability* reflects the degree to which the relative order of individuals on a given construct is maintained across time. It demonstrates individuals' change in rank order over time, that is, the stability of interindividual differences (Robins & Trzesniewski, 2005). *Mean-level change* reflects differences in the average value of a given construct in a specific group across time; it indicates how a group develops on average. Mean-level change can show different *shapes*, that is, linear or nonlinear (e.g., quadratic, cubic). While rank-order stability and mean-level change entail the entire sample (i.e., stability and change on a sample level), *interindividual differences in intraindividual change* focuses on variability in individual development, that is, on the individual level. In addition to the average change on the sample level, some individuals may show an increase, others show a decrease, and some show no change at all.

3.2.2 Theoretical assumptions about self-concept development in adolescence

The development of self-concept facets in early, middle, and late adolescence is the topic of different developmental theories and models (e.g., Harter, 2012b). Especially, the influence of cognitive development stages, contextual changes (e.g., the transition from elementary school to secondary school), and physical changes (e.g., puberty) and their interplay have been considered. In the following, assumptions that could be derived from multiple perspectives on different conceptualizations of development of the self-concept facets targeted in the present study are described.

Different self-concept researchers (e.g., Marsh, 1989; Shavelson et al., 1976; Shavelson & Bolus, 1982) posit that self-concept facets should become more stable regarding rank-order with age.

Further, Shavelson et al. (1976) assume that global self-concepts (e.g., self-esteem) should be more stable than specific self-concepts (e.g., academic and social).

Further, taking a developmental perspective, Shavelson et al. (1976) and Marsh (1989) posit that self-concepts should not only become more stable, but should also become increasingly realistic with age through the continuing integration of feedback from significant others leading to a decline in mean-level of most self-concept facets from childhood to middle adolescence. In stage-environment-fit theory (see Eccles et al., 1993), which focuses on the influence of the context on the development (i.e., the fit of context and the needs of adolescents), it is posited that motivational constructs including academic self-concept and self-esteem in adolescence may decrease in mean-level because of changing contexts (e.g., transition from elementary to secondary school) and environments that do not show a good match with individuals' needs in adolescence. Because not all individuals show the same reaction to the transitions and the same fit to contexts due to, for example, different personalities (see e.g., Block & Robins, 1993), interindividual differences in intraindividual developmental trajectories in multiple self-concept facets can be expected in early adolescence (Cole et al., 2001) and beyond. Another factor that contributes to different developmental trajectories in self-concept facets is that individuals vary in the age in which they enter development stages to a certain degree (e.g., onset of puberty) (e.g., Schaffhuser, Allemand, & Schwarz, 2016)

In middle to late childhood (ages 8-10), comparative assessment with peers and social comparisons are mainly used for self-evaluation (Harter, 2012b). Further, Harter (2012) assumes that self-concept started to become more accurate leading to a decline in self-concept facets. Moreover, individuals are now able to perform a global evaluation of themselves (i.e., self-esteem) that can be verbalized. The transition from elementary to secondary school is assumed to be a critical and stressful period due to changing contexts and the physical (the onset of puberty) and social challenges that affect individuals’ self-perceptions (Anderman & Midgley, 1997). The transition usually occurs in the time frame from late childhood to early adolescence.

In early adolescence (ages 11-13), challenges coincide with an increased salience of social skills and social interactions in different contexts in early adolescence and this endures to middle adolescence (see Harter, 1999, 2012b).

Thus, researchers have suggested that early adolescence is a time of *storm and stress* (see e.g., Steinberg, 2008). Marsh (1989) assumes that many self-concept facets decline during early adolescence but recover during late adolescence.

In middle adolescence (ages 14-16), self-concepts are especially formed through social comparisons with significant others in different areas and contexts. Thus, in middle adolescence, individuals are particularly prone to comparative feedback that might contribute to changes in the self-concepts’ stability and level. Regarding stage-environment-fit theory (Eccles et al., 1993), a mismatch between contextual changes and developmental needs is expected in middle adolescence as problematic leading to a decline in self-esteem and academic self-concept in middle adolescence. Factors causing a mismatch, for example, is ability grouping at a time when students are highly sensitive to social comparison information (Eccles et al., 1993).

Finally, in late adolescence (ages 17-19) most individuals have transitioned through puberty, and cognitively, most adolescents have made the transition from concrete to formal operations which enables them to comprehend complex cognitive linkages (see e.g., Harter, 2012b). In this vein, Harter (2012) assumes that social comparisons now have less influence on and the ideal standards of adolescents gain more impact on self-concept formation. Thus, there is an increasing impact of one’s own ideal standards on self-concept formation.

3.2.2.1 Empirical findings for the development of self-concept in adolescence.

In the following, we present empirical findings for self-esteem, academic self-concepts, and social self-concept structured by different conceptualizations of development on the sample level (rank-order stability, mean-level change) and on an individual level (interindividual differences in intra-individual change).

Self-esteem. A meta-analysis of 50 studies indicated that *rank-order stability* in early childhood is low, and increases throughout adolescence and young adulthood (age 9 to 18 r =.39-.61: Trzesniewski et al., 2003). Results on the *mean-level change* in adolescents' self-esteem were inconclusive. Several studies indicated that mean levels were high in childhood, dropped during adolescence, and increased from adolescence throughout adulthood (Chung et al., 2014; Orth & Robins, 2014; Robins et al., 2002; Robins & Trzesniewski, 2005; Soest et al., 2016). However, a few studies found evidence for increasing self-esteem throughout adolescence (Demo, 1992; Erol & Orth, 2011; Soest et al., 2016; Steiger et al., 2014), while other studies found no mean-level change in self-esteem (Birkeland et al., 2012; Kuzucu et al., 2014). In part, these inconsistent findings might be explained by the method of assessment. Huang (2010) found, in his meta-analysis including 130 samples, that the average increase or decrease in adolescence depended on the self-esteem measure that was used. Fewer studies have investigated whether mean-level change showed a *linear or nonlinear shape*. A recent study by Birkeland et al. (2012) indicated a nonlinear shape rather than a linear shape. Several studies found *interindividual differences in the intraindividual developmental* trajectories (Birkeland et al., 2012; Morin, Maïano, Marsh, Nagengast, & Janosz, 2013; Young & Mroczek, 2003). However, the issue of change in self-esteem at the individual level has rarely been addressed in the literature (Robins & Trzesniewski, 2005) .

General academic self-concept. General academic self-concept increases in *rank-order stability* throughout elementary school (Guay et al., 2003). During adolescence, findings point to moderate to high rank-order stability (Cole et al., 2001, 6-month stability grade 5 to grade 11 rs =.67-.85; 2-year stability between grade 7 and 9 r =.42Gogol et al., 2016). Regarding *mean-level change,* several studies indicated a decrease in general academic self-concept throughout adolescence (Archambault et al., 2010; Fraine et al., 2007; Fredricks & Eccles, 2002; Rieger et al., 2017). With respect to the *shape of mean-level change,* most of the studies that investigated different shapes of mean-level change found a nonlinear shape (see e.g., Birkeland et al., 2012; Cole et al., 2001). Further, the findings of several studies demonstrated *interindividual differences in intraindividual change* of academic self-concept (Archambault et al., 2010; Kuzucu et al., 2014; Young & Mroczek, 2003).

Social self-concepts. Previous studies found high *rank-order stability* of social self-concept of acceptance throughout adolescence (grade 5 to grade 11 rs =.64-.85:Cole et al., 2001; 4-month stability in grade 7 rs =.50-63: Trautwein, 2003). Trautwein (2003) found that rank-order stability in social self-concept of assertion was moderate to high (4-month stability in grade 7; *r*s = .44-.51). Studies investigating *mean-level change* of social self-concept of acceptance indicated an increase throughout adolescence (Cantin & Boivin, 2004; Young & Mroczek, 2003). Harris et al. (2017) found that the social self-concept (opposite-sex) increased. However, Bolognini et al. (1996) found a decrease in their investigation of individuals at the age of 12 to14 years. Trautwein (2003) found an increase in the social self-concept of assertion throughout grade 7 in a sample of German students. Considering different *shapes of mean-level change*, previous research on the social self-concept of acceptance indicated that the shape of change might be nonlinear in nature. For example, Cole et al. (2001) showed that the shape of mean-level changes throughout adolescence are nonlinear. Marsh (1989) found a nonlinear shape of mean-level change in social self-concept (same sex) but not in the social self-concept (opposite sex) throughout adolescence. To our knowledge, no study has yet investigated the shape of mean-level change in social self-concept of assertion. Young and Mroczek (2003) found *interindividual differences in intraindividual development* in social self-concept of acceptance. However, to our knowledge, no studies have yet investigated interindividual differences in intraindividual change for the social self-concepts of assertion.

To sum up, there is consistent evidence for a moderate to high rank-order stability in self-esteem, general academic self-concepts, and social self-concepts. However, previous results on mean-level change are inconclusive. Further, most previous studies did not model the shape of change, that is, if the change displays a linear and nonlinear (e.g., quadratic) trajectory but modeled only a linear form (see e.g., Kuzucu et al., 2014; Young & Mroczek, 2003). However, findings from recent studies point to a nonlinear shape of the mean-level changes in adolescents' general academic self-concept and social self-concept of acceptance. However, no study to date has investigated the shape of the mean-level change of in the social self-concept of assertion. Regarding change on the individual level, several studies found evidence for interindividual differences in intraindividual change in self-esteem, general academic self-concept, and the social self-concept of acceptance. However, to the best of our knowledge, no study investigated interindividual differences in intra-individual development for the social self-concept of assertion.

3.2.2.2 Cross-sectional and longitudinal relations of self-esteem, academic self-concept, and social self-concepts: Theory and findings

Across adolescence, the concurrent correlations between self-esteem and general academic self-concept or social self-concept of acceptance are positive (see e.g., Byrne & Gavin, 1996; Byrne & Shavelson, 1996; Marsh et al., 1988; Tetzner et al., 2016; Wagner & Valtin, 2004). Academic and social self-concepts showed small to moderate positive correlations (e.g., Byrne, 1996). Social self-concept of acceptance and social self-concept of assertion were highly positively correlated in grade 7 and in grade 10 (see Trautwein, 2003). However, how self-esteem and self-concepts are causally related across time still is an open question (Marsh & Yeung, 1998; Wagner & Valtin, 2004). A matter of debate is whether general self-concept (e.g., self-esteem) is predominately determined by specific self-concepts or vice versa. Moreover, whether specific self-concepts are determined by other specific self-concepts is hardly discussed and investigated (Trautwein, 2003).

Bottom-up, top-down, and reciprocal effects. *Bottom-up* effects imply that the direction of longitudinal relation goes from specific self-concepts to global self-concept (e.g., from social self-concept of acceptance to self-esteem). The bottom-up perspective is emphasized in the Shavelson model (1976). According to this model, individuals' self-concepts are formed through feedback from significant others and social interactions with the environment. Thus, the causal flow in the self-concept hierarchy is from the bottom to the top so that, for example, changes in academic self-concept cause changes in self-esteem. However, Shavelson et al. (1976) do not preclude the existence of top-down effect. Therefore, Marsh and Yeung (1998) interpreted the assumption implied in the Shavelson model to be *reciprocal* in nature. Reciprocal relations mean that bottom-up and top-down effects can be found at the same time.

Similar predictions were made by Harter (1999, 2012b), who assumes that self-esteem results from one's accomplishments in important domains (e.g., in the academic domain). Further, feedback obtained from significant others is supposed to be an important determinant of self-esteem. However, Harter (1999) also acknowledged that it is generally possible, to a small degree, that self-esteem has an influence on domain-specific self-concepts.

Contrary to Harter (1999) and Shavelson et al. (1976), Brown (Brown, 1993; Dutton & Brown, 1997) assumed that the causal flow mainly goes from self-esteem to specific self-concepts, that is, *top-down*. Brown (1993) argued that "through transfer of affect processes and halo effects, positive feelings toward the self in general color people's evaluations of their specific attributes" (p. 31). However, he also acknowledged that extreme failure or success in a domain of life may lead to a change in self-esteem, which implies a bottom-up effect. Brown (1993) found evidence for his assumptions by investigating mainly college students using a quasi-experimental design. However, only longitudinal studies allow researchers to draw conclusions about the directionality of relations between self-esteem and specific self-concepts (Marsh & Yeung, 1998). Some self-concept researchers (see Marsh & Yeung, 1998) assume that there is no exchange between global and specific self-concepts or exchange only to a very small degree. Marsh and colleagues (Marsh, 1993b; Marsh & Yeung, 1998) assume that specific self-concepts have little predictive value for the development of self-esteem.

To conclude, the different assumptions are partly opposite of each other and partly contradicting. Thus, the pattern of bottom-up or top-down effects or reciprocal effects remains unclear. In this regard, Marsh and Yeung (1998) stated that "theoretical accounts of the direction of causal flow, however, have been ambiguous, have not adequately operationalized terms, and have not provided any defensible methodological approaches" (p. 511). Suls (1993) concluded, "both Brown and Marsh, who cite strong support for their viewpoints, cannot be right; or, at minimum, a new integrative theory is needed to reconcile the two approaches" (p. X).

Few studies investigated the longitudinal relations between different self-concept facets. In the following results from studies were described which investigated the same or related self-concept facets or a subset of those that were investigated in the present study.

For example, Trautwein (2003; study 2) studied the longitudinal relations between self-esteem, academic self-concept in math and German, and the social self-concepts of acceptance and assertion across three time points in a sample of German 7th graders. He found some evidence for bottom-up effects with regard to math academic self-concept and language academic self-concept on self-esteem (range β = .07-.16). Further, he found weak evidence for top-down effects of self-esteem on specific self-concepts. Bottom-up effects were only observed in relation to general academic self-concepts by Skaalvik and Hagtvet (1990). The authors found that academic self-concept had a small, positive effect on self-esteem in students in grades 6 and 7 (β =.19), but not in a younger cohort in grades 3 and 4. They found also top down-effects between self-esteem and academic self-concept in students in grades 6 and 7. This indicated a reciprocal relation. Wagner and Valtin (2004) studied German students in grades 8 and 9 and assessed self-esteem, general academic self-concept, social self-concepts, and self-concept of physical appearance. They found evidence for top-down effects regarding self-concept of ability (male β =.14; female β =.14), peer acceptance (male β =.43; female β =.31), and physical appearance (male β =.28; female β =.29). Thus, the strongest top-down effect was found from self-esteem at grade 8 to self-concept of peer acceptance at grade 9.

Thus, to sum up, previous studies are inconclusive regarding the direction and the pattern of causal longitudinal (bottom-up, top-down, reciprocal) relations between self-esteem, academic self-concepts, and the two social self-concepts. Overall, effect sizes of significant relations between self-concepts were small in most of the studies (but see Wagner & Valtin, 2004).

Transdimensional effects. Effects between specific self-concepts at the same level or different levels of specificity are called *transdimensional* (Trautwein, 2003). The causal longitudinal relations between specific self-concepts can be unidirectional or reciprocal. Transdimensional effects are frequently investigated with respect to the influence of dimensional comparison processes. According to Möller and Marsh (2013), individuals tend to compare attributes of a domain with attributes of another domain. These comparison processes influence self-concept formation particularly in the academic section of self-concept (I/E model; Marsh, 1986). However, transdimensional effects between academic and social self-concepts and within social self-concepts were scarcely discussed and investigated (Trautwein, 2003). In one of the few studies examining these effects, Trautwein (2003) found weak evidence for positive longitudinal transdimensional effects between social self-concepts of assertion and academic self-concepts in math and the native language in 7th graders. However, he did not investigate the longitudinal relation between both social self-concepts.

Preckel et al. (2013)[15]investigated German students in grade 5 and grade 8 and found that earlier social self-concept of acceptance negatively predicted changes in general academic self-concept over time while earlier social self-concept of assertion positively predicted changes in academic self-concept. Further, social self-concepts were not longitudinally related to each other.

To conclude, results from previous research are inconclusive regarding the pattern of causal longitudinal relations between self-esteem and general academic self-concept, and the two social self-concepts and effect sizes of relations between self-concepts were, in most of the studies, small. Transdimensional effects between general academic self-concept and the two social self-concepts or within social self-concepts were only rarely investigated (Trautwein, 2003). Moreover, it is difficult to compare the existent studies because study designs varied in timespans and in the self-concepts that were taken into consideration. Thus, conclusions on the causal flow between self-concepts throughout adolescence remain elusive.

[15] Preckel et al. (2013) used a partly overlapping and a partly extended sample of the sample included in the present study. Our sample differed in two ways from that of Preckel et al. (2013). First, two cohorts of students from one school in the sample were excluded because this school changed the length of schooling from nine years of secondary education (after grade 4) to eight. Second, we included two additional measurement points in grade 5, and because the sample de-rive from an ongoing project, one assessment wave in grade 8 and data from grade 11 were added.

3.3 Research aims and hypotheses

In a sample of German students from grade 5 to grade 11, comprising the time of late childhood to late adolescence, we investigated the development of self-esteem, general academic self-concept, and the social self-concepts of acceptance and assertion.

The aims of the present study were twofold:

1) We investigated central aspects of change in self-concepts during adolescence on the sample level (i.e., rank-order stability, mean-level change, and the shape of the mean-level change) and on the individual level (i.e., interindividual differences in intraindividual change). Based on previous research (Cole et al., 2001; Trautwein, 2003; Trzesniewski et al., 2003) we expected a moderate to high rank-order stability for all constructs under study. We assume, in line with assumptions of several self-concept researchers (see Harter, 1999, 2012; Marsh, 1989; Shavelson et al., 1976; Shavelson & Bolus, 1982), that multiple self-concept facets should become more stable in rank-order throughout adolescence on the sample level.

Further, regarding the stage-environment-fit theory (Eccles et al., 1993), we expected a negative mean-level change in self-esteem and general academic self-concept specifically in early adolescence. Regarding social self-concept of acceptance, we expected a positive mean-level change in line with the majority of previous research findings (e.g., Young & Mroczek, 2003). We expected that the shape of mean-level change in self-esteem, academic self-concept, and the social self-concept of acceptance follows a nonlinear pattern, in line with previous research (see, e.g., Birkeland et al., 2012; Cole et al., 2001). We take an explorative approach in our examination of the social self-concept of assertion because this self-concept variable has not been studied up to now.

2) We investigated the causal flow between self-concepts. Different theoretical perspectives on the causal longitudinal relations have been posited (see, e.g., Marsh & Yeung, 1998). However, only few studies investigated the causal longitudinal relations, and results were inconclusive (Marsh & Yeung, 1998; Shavelson & Bolus, 1982; Trautwein, 2003; Wagner & Valtin, 2004). Therefore, we did not formulate specific hypotheses regarding the pattern and predominance for *bottom-up*, *top-down*, and *reciprocal* effects. In line with the majority of findings from previous studies, we expected to find small effect sizes for relations between self-concepts facets.

In our examination of *transdimensional* effects, we aimed to replicate and extend previous findings from Preckel et al.'s (2013) study by using an extended sample with more measurement time points across a longer time span. Moreover, we included self-esteem in our analysis. Specifically, as shown by Preckel et al. (2013), we assume that between the middle of grade 5 to the middle of grade 8, students' earlier general academic self-concept should not be related longitudinally to their later social self-concepts of acceptance and assertion. Earlier social self-concept of assertion should be positively related to later general academic self-concept whereas earlier social self-concept of acceptance might be negatively related to later general academic self-concept. Further, social self-concepts might not be causally related between the middle of grade 5 to the middle of grade 8.

Finally, we conduct an explorative analysis of the causal relation for the time span from the beginning of to the middle of grade 5 as well as the time span from grade 8 to grade 11 because there are no findings on these relationships available from previous studies.

3.4 Method

3.4.1 Sample and procedure

Data stem from an ongoing longitudinal study in which students in two federal states of Germany (Rhineland-Palatinate, Bavaria) and 5 schools and 35 classes of the highest track of the German secondary school system (i.e., Gymnasium) participated. The study commenced in 2005, and students were assessed three times in grade 5. This frequency of assessment enabled us to capture the rapid developmental and contextual change in early adolescence after the transition from elementary to secondary school in an adequate fashion.

Further, students were assessed one time in grade 6, grade 8, and grade 11. In the present investigation we analyzed the data from N =1,163 students (46.7% female) who were assessed one week after the beginning of grade 5 (t1) (age: M = 10.54, SD = .43), 11 weeks after the beginning of grade 5 (t2) (age: M = 10.70, SD = .43), in the middle of grade 5 (t3) (age: M = 10.99, SD = .43), in the middle of grade 6 (t4) (age: M = 12.02, SD = .43), in the middle of grade 8 (t5) (age: M = 14.00, SD = .44), and in the middle of grade 11 (t6) (age: M = 16.98, SD = .39). The questionnaire was administered during regular class sessions by trained research assistants using a highly standardized procedure. Student participation was anonymous and voluntary. Written parental consent was obtained for all participating children who were not of legal age.

3.4.1.1 Variables and measures

Self-esteem. Self-esteem was assessed with four items from Rosenberg's (Rosenberg, 1965) Self-Esteem Questionnaire. The items were: "On the whole, I am satisfied with myself," "I feel that I have a number of good qualities," "I am able to do things as well as most other people," and "I take a positive attitude toward myself."

General academic self-concept. General academic self-concept was assessed with a German translation (Kunter et al., 2002) of three items of the short version of the Self-Description Questionnaire (SDQ-II; Marsh, 1990a: "I learn things quickly in most school subjects," "I do well in tests in most school subjects," and "I am good at most school subjects."). Thus, we considered the cognitive part of academic self-concept which shows higher correlations with self-esteem than the affective part (Marsh & Ayotte, 2003).

Social self-concepts. To assess *social self-concept of acceptance* and *social self-concept of assertion* we used three-item short scales developed by (Fend & Prester, 1986). These scales were used in several national studies in Germany (Jonkmann, Trautwein, & Lüdtke, 2009; Trautwein et al., 2004). For the assessment of social self-concept of acceptance we used the items: "Others frequently ignore me when they do something together during breaks," "It does not matter what I do – I am not well received by my peers," and "In class, I sometimes feel like an outsider." The items for *social self-concept of assertion* were: "Sometimes I don't say anything although I am right," "I think I am not as able to take hold of something as others," and "Even if I am right I don't dare to complain about something" (all item translations into English by the authors).

Students responded to all items measuring self-concepts on a 5-point rating scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Items assessing *social self-concepts* were inverted, so that higher scores on all scales indicated higher self-esteem or self-concepts, respectively.

Scale reliability (Cronbach's alpha) of self-esteem and the self-concepts was acceptable to good (Aiken & Groth-Marnat, 2006), ranging from .64 to. 84 (see Appendix A, Table A 1).

3.4.2 Data analysis

All models were estimated with the statistical software Mplus 8 (Muthén & Muthén, 1998-2017). To avoid distorted significance testing due to nested data (i.e., students were nested within classes), we used the “type is complex” option in Mplus. To deal with missing data and nonnormal distributed data we used a robust full information maximum likelihood estimator (MLR). At any time point, 32.5% to 75.5% of the students provided data to estimate variances or covariances (covariance coverage; self-esteem items: 32.5-75.5%; general academic self-concept items: 32.8-75.4%; social self-concept of acceptance items: 38.2-71.3%; social self-concept of assertion items: 32.9-75.2%).

Model fit evaluation was based on the following criteria: Root mean square error of approximation (RMSEA) should below .08 and .05, as guidelines for acceptable and excellent fit, comparative fit indices (CFI) should be greater than .90 and .95, respectively (Hu & Bentler, 1999; Marsh, Hau, & Wen, 2004; Marsh, Hau , & Grayson, 2005). We also reported the chi-square (χ^2) test statistic but due to its sensitivity to sample size, we mainly focused on descriptive fit indices.

To identify and set the scale of the latent variables, we used effect coding (Little, Slegers, & Card, 2006). Therefore, at each time-point and for every construct, we constrained the average of the factor loadings to equal 1.0 and the average of the indicator intercepts to equal 0. Effect coding is advantageous because the latent mean and standard deviation at every time point can be estimated in a meaningful metric, that is, that the latent mean of a construct at a single time point can be interpreted as the average of the items mean weighted by the factor loadings (Little, Preacher, Selig, & Card, 2007). Moreover, regarding measurement invariance, the invariance across time can be tested with respect to the factor loadings and intercepts of all indicator (Little et al., 2006). Further, in all models, the residual variances of corresponding indicators were allowed to correlate across the time points to account for indicator-specific variance across time (correlated uniquenesses see e.g., Bollen & Curran, 2006)

Latent modeling of self-concepts allowed us to test for measurement invariance. To interpret covariance structure (i.e., the magnitude of interindividual differences in intraindividual change across time points), (partial) metric measurement invariance is required; to interpret mean structure (i.e., latent means) across time points, (partial) scalar measurement invariance is required (Little et al., 2007; Vandenberg & Lance, 2000). Therefore, we tested each self-concept separately for measurement invariance. To do this, equality constraints are placed on factor loadings (metric measurement invariance model) and on intercepts (scalar measurement invariance model) of corresponding items across time. We evaluate whether the constraints were justifiable if the change in CFI is smaller than or equal to .01 (Cheung & Rensvold, 2002).

3.4.2.1 Investigating stability and change

To evaluate rank-order stability we inspected latent correlations between latent variables obtained from the scalar measurement model for each self-concept modeled separately. For the evaluation of mean-level changes, the shape of mean-level changes, and interindividual differences in intra-individual change, we specified second-order latent growth curves (Geiser, Keller, & Lockhart, 2013) for each self-concept separately. We modeled both linear and nonlinear (quadratic) change and compared the model fit of the two models using the χ^2 difference test. Specifically, unequal spacing of measurement points across time was considered by fixing the size of the factor loadings on the linear factor (linear change model) and the linear and quadratic factor (quadratic change model) accordingly. The quadratic change model is depicted in Figure 3. To depict the number of individuals who showed an increase, a decrease, or no change, we illustrated the distribution of the estimated linear slope values for individuals in a histogram. We further depicted bivariate scatterplots with the estimated linear slope values for individuals on the x-axis and quadratic slope values on the y-axis (for those constructs that showed a quadratic change only). Moreover, we presented trajectories of individual change for each self-concept facet for 100 randomly selected individuals using spaghetti plots.

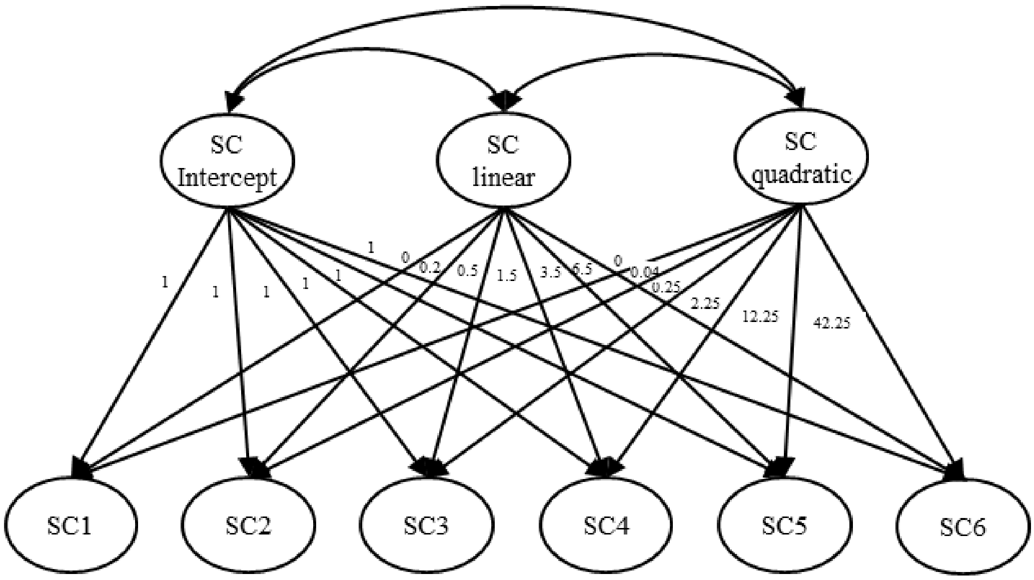


Figure 3. Quadratic second-order latent growth curve model for self-concept (SC) development across six time-points. Latent variables are presented as ellipses. Indicators of latent constructs are omitted. Residuals of corresponding constructs were allowed to correlate.

3.4.2.2 Causal flow between self-concept facets

Cross-lagged panel models are a common method to investigate the directionality of longitudinal relations between variables in correlational longitudinal data[2] (e.g., McArdle & Nesselroade, 2014) and was most used to investigate the causal flow between self-concepts. Thus, we specified a six-wave cross-lagged panel model including all four constructs in which we estimated the stability of the constructs across subsequent measurement waves (self-concept facet → self-concept facet), bottom-up effects (general academic self-concept, social self-concepts → self-esteem), top-down effects (self-esteem → general academic self-concept, social self-concepts), and trans-dimensional effects (general academic self-concept → social self-concepts; social self-concepts → general academic self-concept; social self-concept of acceptance→ social self-concept of assertion; social self-concept of assertion→ social self-concept of acceptance) (see Figure 4).

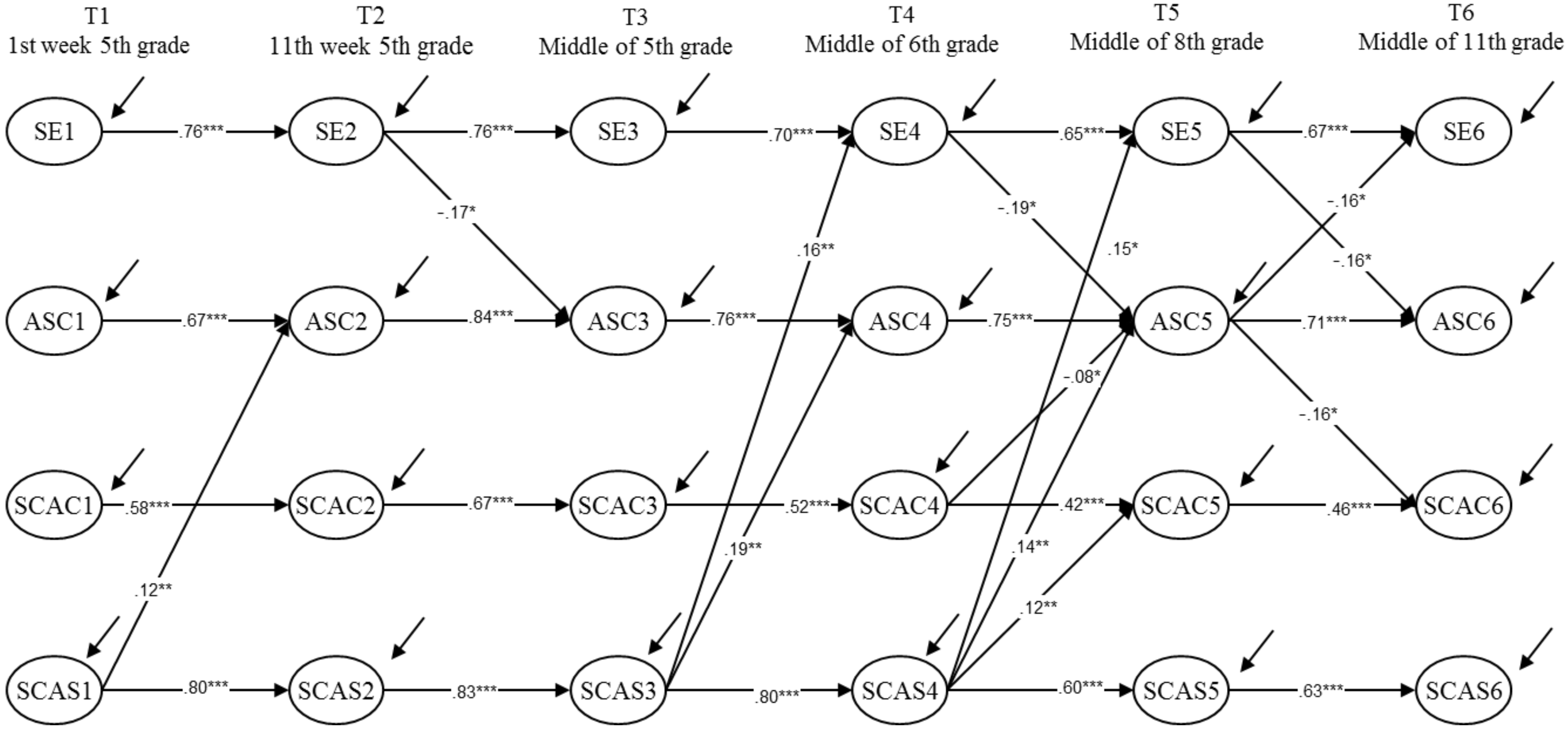


Figure 4. Model 1: Multiwave cross-lagged model between self-esteem (SE), general academic self-concept (ASC), social self-concept of acceptance (SCAC), social self-concept of assertion (SCAS). Standardized model parameters are shown. Only significant paths are shown. Within each wave, the latent constructs are assumed to be correlated. Residuals of corresponding constructs were allowed to correlate. * $p < .05$, ** $p < .01$, *** $p < .001$

N	$\chi 2(df)$	CFI	RMSEA (90%CI)	SRMR
1163	4016.689 (2700)***	.938	.020 (.019—.022)	.061

3.5 Results

3.5.1 Preliminary analyses

Cross-sectional descriptive correlations (see Appendix A, Table A 1) and latent correlations (see Appendix A, Table A 2) showed that self-esteem displayed the highest correlation with general academic self-concept (descriptive: *r*s: 60 -.32; latent: *r*s: .77-.32) followed by the social self-concept of assertion and the social self-concept of acceptance. General academic self-concept and the social self-concept of assertion were moderately to highly correlated (descriptive: *r*s: .13-.23; latent: *r*s: .15-.34); the social self-concept of acceptance and the social self-concept of assertion were uncorrelated or moderately correlated (descriptive: *r*s: .05-.27; latent: *r*s: .02-.37). The social self-concept of assertion and the social self-concept of acceptance were moderately to highly correlated (descriptive: *r*s: .32-.45; latent: *r*s: .40-.61).

Measurement invariance. Full scalar measurement invariance could be assured for general academic self-concept and the social self-concept of assertion, as the ΔCFI values between the metric and scalar measurement invariance models were below the cut-off value of .01. Partial scalar measurement invariance (Dimitrov, 2010) could be confirmed for self-esteem and the social self-concept of acceptance. With regard to self-esteem, three constraints concerning intercepts had to be released and one constraint concerning intercepts for social self-concept of acceptance. All models provided a good model fit (see Appendix A, Table A 3).

3.5.2 Main analysis

Rank-order stability. Rank-order stabilities were moderate to high (self-esteem, social self-concept of assertion: *r*s = .32-.71; social self-concept of acceptance: *r*s = .30-.70; general academic self-concept: rs = .36-.69) (see Table 2). Between adjacent time points there were higher correlations than between nonadjacent time points or subsequent time points that were further apart (between grades 6 and 8 [t4-t5] and grades 8 and 11 [t5-t6]). To compare the stability coefficients across the time within the self-concept facets, stability coefficients were transformed into the same metric (i.e., stability across one year). The annualized stability coefficients indicated growing stability with increasing grade level for all self-concepts (see Table 2).

Table 2. Latent means and standardized latent mean-level changes (*d*) and correlations (rank-order stability) of self-esteem (SE), academic self-concept (ASC), social self-concept of acceptance (SCAC), and social self-concept of assertion (SCAS)

				mean-level change					rank-order stability				
	M	*SE*	*var*	d_{1x}	d_{2x}	d_{3x}	d_{4x}	d_{5x}	r_{1x}	r_{2x}	r_{3x}	r_{4x}	r_{5x}
SE													
T1	4.096	.023	.278										
T2	4.076	.026	.320	-.038					.712*** (.171)				
T3	3.981	.023	.349	-.218***	-.180***				.603***	.710*** (.226)			
T4	3.982	.031	.386	-.216***	-.178***	.002			.492***	.544***	.652*** (.652)		
T5	3.990	.038	.339	-.201**	-.163*	.017	.015		.368***	.388**	.480***	.588*** (.767)	
T6	4.006	.045	.358	-.171	-.133	.047	.046	.030	.324***	.410***	.451***	.440***	.525*** (.807)
ASC													
T1	4.244	.027	.288										
T2	4.015	.029	.349	-.427***					.644*** (.101)				
T3	3.909	.034	.453	-.624***	-.196***				.541***	.694*** (.205)			
T4	3.756	.034	.507	-.909***	-.482***	-.285***			.455***	.576***	.668*** (.668)		
T5	3.553	.040	.612	-1287***	-.861***	-.663***	-.378***		.387***	.467***	.582***	.619*** (.786)	
T6	3.486	.032	.557	-1.412***	-.986***	-.788***	-.503***	-.125	.363***	.375***	.425***	.478***	.619*** (.852)
SCAC													

T1	4.325	.028	.486										
T2	4.359	.033	.527	.049					.608*** (.075)				
T3	4.244	.042	.642	-.116*	-.164*				.503***	.703*** (.217)			
T4	4.330	.044	.647	.007	-.046	.123*			.431***	.420***	.541*** (.541)		
T5	4.409	.035	.567	.120*	.072	.237**	.113		.354***	.301***	.331***	.413*** (.642)	
T6	4.559	.031	.287	.336***	.287***	.452***	.328***	.215***	.296***	.208***	.261***	.402***	.458*** (.771)
SCAS													
T1	3.573	.032	.575										
T2	3.663	.036	.726	.118					.712*** (.171)				
T3	3.552	.037	.753	-.025	-.146**				.678***	.763*** (.310)			
T4	3.767	.039	.634	.256***	.137**	.284***			.600***	.586***	.746*** (.746)		
T5	3.903	.038	.566	.435***	.317***	.463***	.179*		.391***	.400***	.481***	.527*** (.726)	
T6	3.889	.032	.507	.417***	.299***	.444***	.161*	-.018	.319***	.266***	.393***	.402***	.524*** (.806)

Note. d = (mean of Tt = mean of Tt-1)/ standard deviation at T1. T = Time. Column d_{1x}/r_{1x} contains the effect sizes for the mean level difference/correlations depicting the time 1 measure. The annualized rank-order stabilities are depicted in parentheses. Latent correlations are annualized for time interval by taking the nth root of the estimate, where n is equal to the length of the interval (i.e., number of weeks/years between assessments) (see Chung, Hutteman, van Aken, & Denissen, 2017). Mean-level differences were tested with Wald χ^2 test. * $p < .05$. ** $p < .01$. *** $p < .001$.

Mean-level change. Inspection of standardized mean differences (see Table 2) showed no significant mean-level change in self-esteem between consecutive time points except from t3 to t4 (d_{23} = -.128, $p \leq .001$). Regarding general academic self-concept, there were small mean-level changes in all but one consecutive comparison between t5 and t6. Considering social self-concepts, there were small mean-level changes with respect to all but two consecutive comparisons (social self-concept of acceptance: between t1 and t2 and t4 and t5; social self-concept of assertion: between t1 and t2 and t5 and t6).

Table 2 depicts the means of all self-concept facets at the different time points and the effect size in mean-level change (standardized difference in mean-level change; Cohen's *d*)[2][16] between pairwise time lags. Descriptively, mean-level changes from t1 to t6 for general academic self-concept indicated a strong decline, a slight decrease in both social self-concepts, and no change in self-esteem.

Shape of mean-level change. Results and model fit indices of the second-order latent growth curve models are presented in Appendix A, Table A 3. Figure 5 illustrates the observed mean-level change and the estimated mean-level change derived from the linear and quadratic second-order latent growth curve models. Results of the second-order latent growth curve models (see Appendix A, Table A 3) indicated that self-esteem and social self-concept of acceptance showed a linear mean-level change because of the nonsignificant χ^2 difference test between the linear and quadratic model. Self-esteem showed no significant mean-level change. General academic self-concept showed a quadratic trend. General academic self-concept decreased in early to middle adolescence

[16] Effect sizes can be considered as small (d = 0.2), medium (d = 0.5), and large (d = 0.8) (Cohen, 1988)

with a lower rate of change in late adolescence. Social self-concept of acceptance increased linearly throughout adolescence. Social self-concept of assertion showed a nonlinear trend with an increase in early and middle adolescence and a lower rate of change in late adolescence.

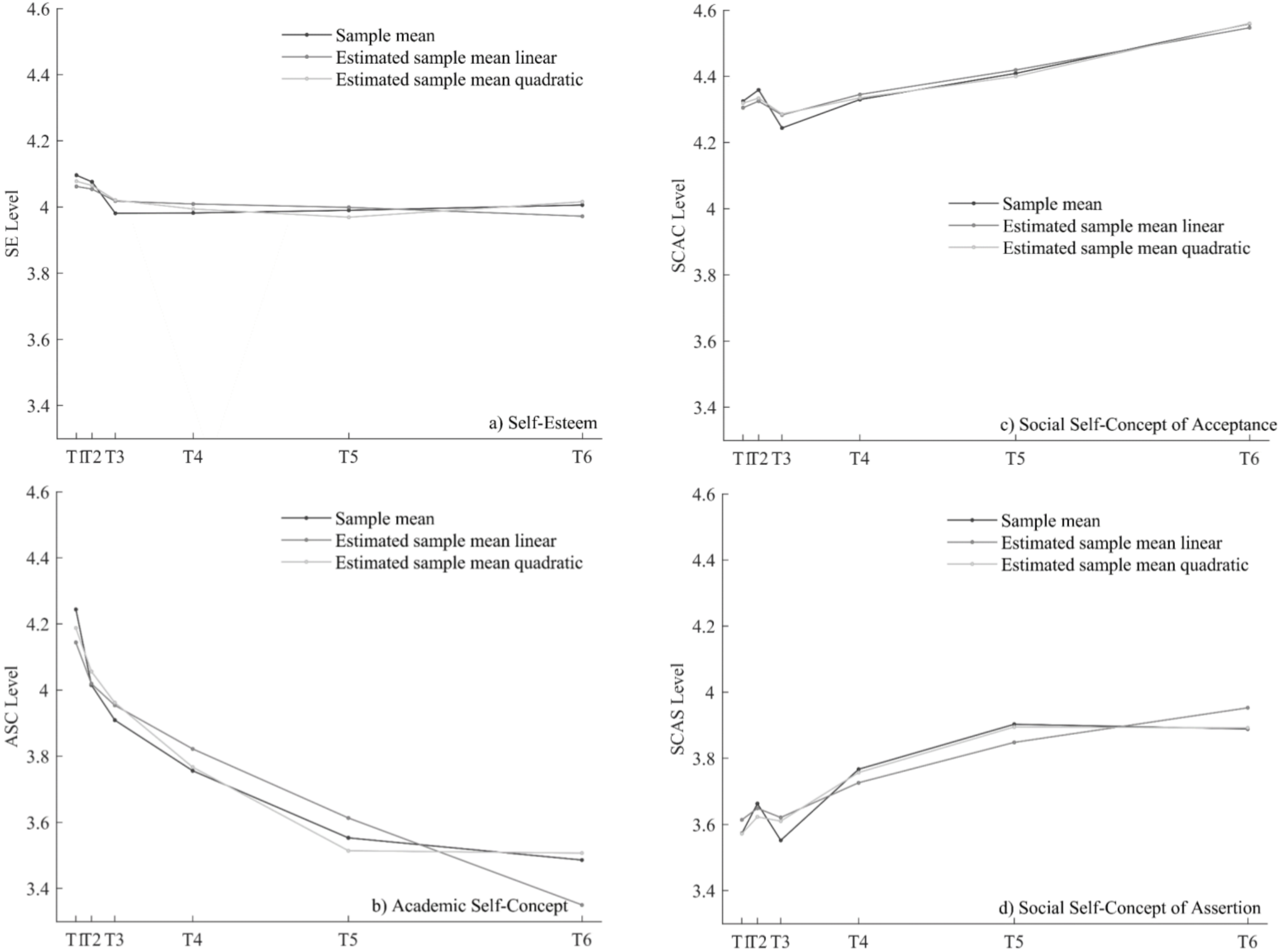


Figure 5. Mean-level changes in latent means derived from the longitudinal CFA models for each construct and estimated latent means derived from the linear and quadratic second-order latent growth curve models. T1 = Time 1; T2 = Time 2; T3 = Time 3; T4 = Time 4; T5 = Time 5; T6 = Time 6.

Self-concepts showed substantial interindividual differences in the initial level at the beginning of grade 5, indicated by a significant intercept factor variance. Further, there were interindividual differences in intraindividual change indicated by a significant slope factor variance. This is illustrated by the distribution of the estimated slope value for each individual (see the histograms; Appendix A, Figure A 1) and the bivariate scatterplots with linear slope estimates at the x-axis and the quadratic slope estimates at the y-axis for general academic self-concept and the social self-concept of assertion (Appendix A, Figure A 2). As depicted in the Appendix A (Figure A 1), 33.50% of the adolescents showed a positive mean level change in self-esteem (estimated linear slope value > .00), 77.60% in the social self-concept of acceptance, and 91.25% in the social self-concept of assertion. Regarding general academic self-concept, the majority of individuals showed a decrease (99.97%). Further, Appendix A, Figure A 3) shows the developmental trajectories of self-concepts of 100 randomly selected individuals of the sample (spaghetti plots).

Causal flow between self-esteem and specific self-concepts. The cross-lagged model provided a good fit (χ^2 (2700) = 4016.589, CFI = .938, RMSEA = .020, SRMR = .061). Results from the cross-lagged model are shown in Figure 4, and in this figure, only significant effects are depicted. All path coefficients are shown in the Appendix A, Table A 4.

Bottom-up effects. There were significant positive effects from the earlier social self-concept of assertion to later self-esteem between grades 5 and 6 (t3) (β = .16) and grades 6 and 8 (t4) (β = .15). That is, students who felt more socially competent at an earlier time point showed higher self-esteem at a later time point than students with a lower social self-concept of assertion.

Top-down effects. Students' earlier self-esteem had a significant negative effect (range β = -.16 to -.19) on later general academic self-concept from week 11 in grade 5 to the middle of grade 5 (t2-t3) and from grade 6 to grade 8 (t4-t5). Further, there was a reciprocal effect between general academic self-concept and self-esteem from grade 8 to grade 11 (t5-t6).

Transdimensional effects. Results indicated positive effects from students' earlier social self-'concept of assertion to their later social self-concept of acceptance from the beginning of grade 5 to 11 weeks later (t1- t2) (β = .13). Moreover, there was a negative effect from earlier general academic self-concept to later social self-concept of assertion from grade 8 to grade 11 (t5-t6) and a negative effect from earlier social self-concept of acceptance to later general academic self-concept from grade 6 to grade 8 (t4-t5) (β = -.08).

3.6 Discussion

In the present study, we investigated the development of four vital self-concept facets in adolescence comprising self-esteem, general academic self-concept, and the social self-concepts of acceptance and assertion in a large sample of German students from grade 5 to grade 11 in a study comprising six measurement waves. Our research aims were twofold: First, we investigated the stability and change of self-concept facets on a sample level and on an individual level. On a sample level, we found that all self-concepts became more stable (rank-order) over time but were different in mean-level change and the shape of mean-level changes. On the individual level, there were remarkable interindividual differences in the development of all self-concept facets. Our second research aim sought to uncover the longitudinal the longitudinal causal relations between self-concepts. We found some longitudinal causal relations between self-concepts that mainly occurred between grades 6 and 8 (t4-t5). The pattern of bottom-up effects and top-down effects was relatively balanced. Only one reciprocal effect was observed. Further, we found evidence for transdimensional effects. Before discussing our findings, we point to the limitations of the present study and derive from these suggestions for future research.

3.6.1 Limitations

Although the sample in the present study was large, participating students are only from the highest track (i.e., Gymnasium) of the German secondary school system. Therefore, our findings should be replicated with a sample comprising other academic tracks or more heterogeneous achievement groups. School tracks not only differ in achievement level but also in learning conditions and the development of ability (Becker, Lüdtke, Trautwein, Köller, & Baumert, 2012). For instance, Thus, it is possible that general academic self-concept and both social self-concepts develop differently in different school types. Indeed, in his investigation of adolescents in 7th and 10th grade in Germany, Trautwein (2003) demonstrated that the social self-concepts of acceptance and assertion both showed a more positive development in adolescents attending the highest school track as compared to a comprehensive school (i.e., the intermediate school type). However, regarding the development of self-esteem, different development trajectories were found between students attending a general school (i.e., the lowest school type) compared to a comprehensive school but not between students attending the highest track school as compared to a comprehensive school.

In addition, we focused on multidimensional self-concepts but we did not investigate the full range of specific self-concepts. Thus, results regarding the causal flow between self-esteem and other self-concept facets may be different to those including social self-concepts and general academic self-concept. For example, Morin, Maïano, Marsh, Janosz, and Nagengast (2011) showed that earlier body image self-concept influenced later self-esteem in adolescence but not vice versa, indicating a bottom-up effect.

In sum, our results are limited to self-esteem, general academic self-concept, and the two social self-concepts with respect to bottom-up, top-down, reciprocal, and transdimensional effects. However, we focused on these self-concepts because major developmental challenges in adolescence concern the academic and the social domain (e.g., Havighurst, 1979)

3.6.1.1 Stability and change of self-concept facets during adolescence

We assume, in line with assumptions of several self-concept researchers (see Harter, 1999, 2012; Marsh, 1989; Shavelson et al., 1976; Shavelson & Bolus, 1982), that multiple self-concept facets should become, on a sample level, more stable in rank-order throughout adolescence. In general, our findings support this assumption. In particular, we found that self-esteem, general academic self-concept, and the social self-concept of acceptance increased in rank-order stability during adolescence. Rank-order stabilities were moderate to high, a finding which indicates that individuals hardly changed their rank in relation to others. This pattern of results was comparable for the social self-concept of assertion, a construct hardly studied in previous research. Moreover, results revealed that descriptively all self-concept facets we investigated were equally stable. This finding contradicts the assumptions of Shavelson et al. (1976) that global self-concept facets (i.e., self-esteem) should be more stable than specific self-concept facets. However, like the majority of previous research, we also found no support for this assumption (e.g., Byrne, 1986; Shavelson & Bolus, 1982)

Focusing on the sample level we confirmed our expectation that self-concept facets would differ in mean-level changes and the shape of mean-level changes. With reference to the assumption of the stage-environment-fit theory (Eccles et al., 1993; Eccles & Roeser, 2009), we expected that self-esteem and general academic self-concept would decrease. As expected, general academic self-concept decreased in its mean level. However, in contrast to expectations, self-esteem showed no change in its mean level.

In line with findings from previous research, we also found positive mean-level changes for the social self-concepts of acceptance and assertion. The shape of mean-level change differed between the self-concepts. General academic self-concept and social self-concept of assertion showed a nonlinear (i.e., quadratic) change while social self-concept of acceptance showed a linear change.

Moreover, we expected that the shape of the mean-level change for self-esteem, general academic self-concept, and the social concept of acceptance is nonlinear. Note that regarding social self-concept of assertion we investigated the shape in an explorative manner because no study had investigated it up to now. In contrast to expectations, self-concept facets differed in their shape of mean-level change. In particular, general academic self-concept and the social self-concept of assertion showed a nonlinear (quadratic) shape whereas the social self-concept of acceptance increased linearly. Thus, in line with theoretical assumptions (e.g., Harter, 2012b; Marsh, 1989; Marsh & Ayotte, 2003) stating that self-concepts should decline in early and middle adolescence and then stabilize, we found a lower rate of mean-level change in general academic self-concept and the social self-concept of assertion in late adolescence. The finding that the two social self-concepts differed in their shape of change might indicate that different processes underlie the development of these social self-concepts.

An explanation might be that there are several ways to gain social acceptance. During adolescence it becomes increasingly possible for individuals to choose their own social niches (e.g., McElhaney et al., 2008). In these they might feel more and more accepted with time. Contrary, as expected in the social skill approaches to social competence (cf., Rose-Krasnor, 1997) being socially competent requires unique competencies consisting of social, emotional, and cognitive skills. Once established they may stay constant. Particularly the time after the transition from elementary to secondary school might be important to establish the perception to have the ability to assert oneself in social interactions. Individuals have to find their place in new social contexts (i.e., classroom). However, the offered explanations are rather speculative. Future studies should examine these processes leading to different shape of mean-level changes in social self-concept of acceptance and assertion.

Furthermore, and as expected, we found that all self-concept facets under study exhibited interindividual differences in intraindividual development. Although on average nearly all students experienced negative development in their general academic self-concept, the decrease differed meaningfully in size but only some students (0.03%) experienced an increase (see Appendix A, Figure A 1). To shed more light on factors that contribute to negative development, it would be very informative for educational practice to investigate to investigate environment-related and person-related predictors to explain this change.

Moreover, investigating interactions between the environment (e.g., teaching practice) and person (e.g., motivational goals) variables might provide information regarding adaptive instruction methodology and course design. In the same vein, some individuals may experience a decline in their social self-concepts. As mentioned, it is important to investigate both environment and person predictors to explain the variability in change. This might contribute to the understanding of for which persons with which particular person characteristics and environment characteristics and combinations of these are more beneficial for fostering the positive development of their social self-concepts.

3.6.1.2 Causal flow between self-concept facets: Bottom-up, top-down, and reciprocal effects.

The longitudinal interplay between the differing self-concepts is a matter of debate (Marsh & Yeung, 1998; Trautwein, 2003; Wagner & Valtin, 2004). Particularly, researchers ask whether changes in self-esteem cause changes in specific self-concepts and/or vice versa. Some researchers assume a predominance of bottom-up processes (Harter, 1999) while others propose top-down processes (Brown, 1993). Finally, the Shavelson model (1976) implies reciprocal effects (Marsh & Yeung, 1998). Furthermore, previous studies regarding this topic are scarce, and findings were mixed.

We found some evidence for bottom-up effects; these effects differed in sign and were of small size. We only observed significant top-down effects with regard to general academic self-concept. From a theoretical point of view, these effects were somewhat perplexing because they were negative in sign, which means that students with lower self-esteem at an earlier time point had a higher general academic self-concept at a later time point and vice versa. In previous studies, however, the top-down effects found for general academic self-concepts were positive in sign (see Trautwein, 2003; Wagner & Valtin, 2004). This finding might in part be explained by sample characteristics. We included only students from the highest track of the German school system; thus, our sample can be classified as a more homogeneous ability group compared to the heterogeneous samples used in previous studies (Trautwein, 2003; Wagner & Valtin, 2004). In particular, the ability group is the one with the highest scholastic requirement of the German school system. Through the control of the influence of earlier general academic self-concepts to self-esteem at a later time point in the cross-lagged panel model, self-esteem at a later time point can be expected as the unique (nonacademic) part of self-esteem. This unique (nonacademic) part of self-esteem, according to Skaalvik and Hagtvet (1990),

> "may be associated with accumulated nonacademic experiences over time, such as sports and other leisure-time activities. Such activities may be time consuming and may attract attention away from school work and educational demands. Emphasizing nonacademic activities, which may particularly be the case for students who enhance their self-esteem by such activities, could therefore contribute to deteriorated academic performance" (p. 306).

Given that achievement positively causally predicts general academic self-concept longitudinally (see e.g., Guay et al., 2003) and that nonscholastic activities (i.e., leisure activities) are more likely to conflict with scholastic performance in this group of individuals because they belong to the track with the highest scholastic requirements in the German school system, the negative longitudinal relations between self-esteem and general academic self-concept might be explained. However, this explanation is rather speculative and requires future research. A further explanation for these contradictory results may lie in the chosen statistical method. In line with the majority of previous research, we investigated the causal relations between self-concepts with cross-lagged panel models. However, within the method of cross-lagged panel analysis, there is no explicit reference to theory of change. In particular, change on the sample level is not distinguished from the change on an individual level. If stability of the constructs studied is caused by time-invariant traits (e.g., self-esteem), the inclusion of autoregressive parameters might have failed to adequately control for the influence of stability (see the critic of cross-lagged models by Hamaker, Kuiper, & Grasman, 2015). This also implies that different trait-like stabilities of constructs across studies may affect the cross-lagged paths, which may even change in sign. Indeed, previous research indicated different mean-level changes when taking different school tracks into account (see Trautwein, 2003). Thus, the chosen statistical method probably may have contributed to the perplexing top-down effects of self-esteem on general academic self-concept. There are several approaches to control for a trait-like stability (Hamaker et al., 2015). Thus, it is a promising avenue for future research to compare the effects using different statistical methods for investigating the longitudinal causal relations between self-esteem and specific self-concepts. However, approaches to control for a trait-like stability are generally more restrictive, for example, they assume stationarity which might be an unrealistic assumption especially if investigating long time spans (cf., Hamaker et al., 2015).

Note, in this regard, that previous studies have primarily used cross-lagged panel models, and these indicated that a student's earlier social self-concept of assertion positively predicted the student's self-esteem from grade 5 to grade 6 (t3-t4) and from grade 6 to grade 8 (t4-t5). That is, students who feel more socially competent at an earlier time point tended to have higher self-esteem at a later time point. These results are plausible, particularly for the time of early adolescence when individuals are confronted with major social challenges (e.g., maintain positive peer relations; Harter, 2012b). For this purpose, adolescents need to feel that they are socially competent in social interactions (Wentzel & Erdley, 1993).

However, whereas the earlier social self-concept of assertion positively predicted later self-esteem, no effects from students' earlier social self-concept of acceptance to later self-esteem were observed. It might be concluded from this result that for self-esteem it seems to be more important for adolescents to perceive themselves as being socially competent than to be socially accepted by their peers. An explanation for this finding is that social skills such as social competence can be seen as a general life skill (Elias, 1995). Thus, feelings of social competence should also affect self-esteem as a more global construct. Moreover, social skills are not absolutely necessary to perceive oneself as socially accepted. It is also possible that individuals might think they are socially accepted because of, for example, their physical attractiveness (Salvia, Sheare, & Algozzine, 1976).

Overall, in line with previous research, we found weak evidence for causal longitudinal relations between self-esteem and specific self-concept facets. To answer the question of the direction of causal flow, researchers should pay attention to underlying processes. One may consider that strategies of self-enhancement or devaluation of different areas suppress causal effects between self-esteem and specific self-concepts. Thus, for example, although general academic self-concept and self-esteem changed in our longitudinal study, changes in general academic self-concept may not be predictive for changes in self-esteem because individuals devaluate the individual importance of general academic self-concept as a self-protective strategy. This perspective is in line with Hattie (2004), who posited the need for investigating strategies that individuals use to maintain and protect their own self-esteem. Therefore, for future research it is a promising avenue to test longitudinal causal relations between self-esteem and specific self-concepts after controlling for the effect of strategies of self-enhancement in order to protect self-esteem.

3.6.1.3 Transdimensional effects

Previous research paid little attention to the longitudinal causal relations between academic and social self-concepts (see Trautwein, 2003). Regarding the investigated transdimensional effects in our study, it should be noted that it is a replication of the results of the Preckel et al. (2013) study with an extended study design comprising more time points and a longer time span. Moreover, self-esteem was additionally included in the present study. Preckel et al. (2013) investigated the causal interplay between the general academic self-concept, social self-concept of acceptance and assertion in grade 5 and grade 8 in a study comprising three measurement points. In the present study, we found some evidence for transdimensional effects between academic and social self-concepts. Specifically, replicating Preckel et al.'s (2013) findings for adolescents from grade 5 to grade 6 (t3-t4) and grade 6 to grade 8 (t4-t5), we found positive effects of earlier social self-concept of assertion on later general academic self-concept. This effect may be explained by the fact that students with better social skills tend to feel more confident about demonstrating their academic skills in the classroom. This feeling of confidence leads to more participatory behavior of the students that, in turn, leads to better grades. Here it is important to note that higher grades are reciprocally and positively related to general academic self-concept (Marsh & O'Mara, 2008). Earlier social self-concept of acceptance was negatively related to later general academic self-concept from grade 6 to grade 8 (t4-t5).

Although the social self-concepts of acceptance and assertion were concurrently moderately to highly correlated longitudinally, in the present study only prior social self-concept of assertion positively influenced later social self-concept of acceptance from grade 6 to grade 8 (t4-t5). This may be because social acceptance can be a result of being socially skilled but social assertion might not be a necessary requirement. That is, individuals can feel socially accepted for different reasons. For example, individuals who were evaluated by peers as physically attractive were more accepted by peers (Salvia et al., 1976).

Replicating the findings by Preckel et al. (2013), we found that earlier social self-concepts of acceptance and assertion are not causally related longitudinally to later general academic self-concept for the time span of grade 5 to grade 8. Beyond that, we found that general academic self-concept negatively predicted changes in social self-concept of acceptance from grade 8 to grade 11 (t5-t6). However, the influence was close to zero.

3.6.1.4 Practical implications

Fostering a positive development of students' self-concept is supposed to be a major educational goal (e.g., Marsh & Hau, 2003a). On the one hand, knowledge on the longitudinal causal relations between self-concepts provides information if fostering one self-concept facet can have side effects on the formation of other self-concept facets.

Further, investigating the development of self-concept facets on a sample level as well as on an individual level gives valuable hints as to whether self-concept interventions should be implemented for all students simultaneously or only for a selected group of students. Further, it is important to know the most likely developmental time frame for changes (e.g., early, middle, or late adolescence).

We found that general academic self-concept decreased for nearly all students and most strongly after the transition to secondary school, that is, in early adolescence. Thus, interventions which foster general academic self-concept should be implemented early in adolescence targeting all students. Although both social self-concepts increased throughout adolescence, there was a small number of students who showed a decrease throughout adolescence. For these students, programs that aim to strengthen social competence (e.g., Christopher, Nangle, & Hansen, 1993; Nangle, Hansen, Erdley, & Norton, 2009) may help to provide a positive social self-concept development. The importance of a positive social self-concept development is supported by our findings of positive effects from a student's former social self-concept of assertion to his or her later general academic self-concepts at the beginning of grade 5 to grade 8 (t1-t5) and self-esteem. Implementing interventions right at the beginning of grade 5 that aim to promote social skills and, therefore, social self-concept development may have positive side-effects considering general academic self-concept and self-esteem. Hence, implementing early interventions promoting social skills is particularly promising in a number of ways—by fostering a positive development of social self-concepts, general academic self-concepts, as well as self-esteem in adolescents.

4 Study 2

Profile formation of academic self-concept in elementary school students in grades 1 to 4[17]

[17] Schmidt, I., Brunner, M., Keller, L., Scherrer, V., Wollschläger, R., Baudson, T. G., & Preckel, F. (2017). Profile formation of academic self-concept in elementary school students in grades 1 to 4. PloS one, 12(5), e0177854. https://doi.org/10.1371/journal.pone.0177854

4.1 Abstract

Academic self-concept (ASC) is comprised of individual perceptions of one's own academic ability. In a cross-sectional quasi-representative sample of 3,779 German elementary school children in grades 1 to 4, we investigated (a) the structure of ASC, (b) ASC profile formation, an aspect of differentiation that is reflected in lower correlations between domain specific ASCs with increasing grade level, (c) the impact of (internal) dimensional comparisons of one's own ability in different school subjects for profile formation of ASC, and (d) the role played by differences in school grades between subjects for these dimensional comparisons. The nested Marsh/Shavelson model, with general ASC at the apex and math, writing, and reading ASC as specific factors nested under general ASC fitted the data at all grade levels. A first-order factor model with math, writing, reading, and general ASCs as correlated factors provided a good fit, too. ASC profile formation became apparent during the first two to three years of school. Dimensional comparisons across subjects contributed to ASC profile formation. School grades enhanced these comparisons, especially when achievement profiles were uneven. In part, findings depended on the assumed structural model of ASCs. Implications for further research are discussed with special regard to factors influencing and moderating dimensional comparisons.

4.2 Theoretical background

Academic self-concept (ASC) is comprised of mental representations of one's abilities in academic subjects. ASC plays an important role in educational psychology because it influences scholastic achievement (e.g., Marsh & Martin, 2011), academic motivation and affect (Craven & Marsh, 2008), and educational choices (Dickhäuser, Reuter, & Hilling, 2005). Understanding the development of ASC across the school career and its determinants is therefore an important issue.

The investigation of ASC is intrinsically tied to theories about its structure. Research on the structure of ASC demonstrated that, from age 4 onward, ASC is a multidimensional construct with separate, domain-specific mental representations of ASCs for various achievement domains (such as specific school subjects) (e.g., Marsh, 1990b; Marsh, Ellis, & Craven, 2002; Marsh & Hocevar, 1985) a finding that generalizes across countries (e.g., Marsh & Hau, 2004; Marsh, Hau, Artelt, Baumert, & Peschar, 2006). Further, subject-specific ASC is comprised of a cognitive and an affective component (e.g., cognitive: "Math is one of my best subjects"; affective: "I like math"). However, in line with most research on ASC (e.g., Brunner et al., 2010), the present study focuses on the cognitive component. Besides the multidimensionality of ASC, there is evidence for a general ASC, which can be assessed by items such as "I am a good student." Nonetheless, exactly how general ASC should be incorporated into the structural model of ASC and at which hierarchical level remains a matter of debate. This is crucial for research questions focusing on a key aspect of the differentiation of ASC, that is, the formation of the profile of one's own strengths and weaknesses. For example, one student may have a higher general ASC than another student. Still, both students may think they are better in math than in their native language. That is, despite mean level differences, both students can share a similar ASC profile. The most recent structural model of ASC, the nested Marsh/Shavelson model (NMS model) (Brunner et al., 2008; see also Brunner et al., 2009; Brunner et al., 2010), seems especially appropriate for tackling questions on the profile formation of ASC because it allows to control for the general ASC level in students' profiles, thus directly depicting students' ASC profiles of subjective strengths and weaknesses in particular subjects. Further, the NMS model accounts for the shape of students' profiles. There is broad empirical support for the NMS model of ASC in secondary school students (Brunner et al., 2008; Brunner et al., 2009; Brunner et al., 2010). However, to our knowledge, no studies have yet examined how well the NMS model of ASC captures the different facets of ASC in elementary school students. To understand ASC profile formation, it is important to take different sources of information into account. It is well documented that ASC are formed on the basis of multiple sources, such as feedback from significant others, and against multiple frame of references (Skaalvik & Skaalvik, 2002). The internal/external frame of reference model (I/E model; Marsh, 1986) posits that students mainly use an external (social) and an internal (dimensional) frame of reference to evaluate their subject-specific ASC. Within a student's dimensional frame of reference, the student compares his or her achievements across subjects, and this is assumed to be a central source for ASC profile formation (Möller & Marsh, 2013). Most support for the I/E model stems from secondary school students (Möller et al., 2009); a few studies also found support for it in samples of elementary school students from grade 3 onwards (e.g., Ehm, Lindberg, & Hasselhorn, 2014; Möller & Marsh, 2013; Pinxten et al., 2015). However, none of these studies used the NMS model of ASC although, as outlined above, it is particularly well suited for investigating the impact of dimensional comparisons on profile formation. Moreover, research on variables that moderate the effect of dimensional comparisons is still rare (Möller & Marsh, 2013). Rost, Dickhäuser, Sparfeldt, and Schilling (2004; 2005) identified the difference between math and native language grades as a moderator. However, Rost et al. drew on data from secondary school students only; it therefore remains unclear to what extent these findings also apply to elementary school students. In sum, whereas the structural models of ASC and the factors influencing ASC profile formation are well studied in adolescents, research on elementary school students is comparably scarce. Expanding our knowledge to younger students is crucial, for example, to understand the development of ASC in this age group and to implement early interventions to support the positive development of ASC. The purpose of this article is to fill this gap. In a large quasi-representative cross-sectional sample of German elementary school students in grades 1 to 4, we investigated (a) the profile formation of ASC using alternative structural models of ASC; (b) the impact of external and dimensional comparisons within the I/E model; and (c) the moderating effect of differences in teacher-assigned grades in the subjects math and German on the strength of dimensional comparisons. To this end, we applied the NMS model and alternative structural models of ASC and systematically compared the results for these models.

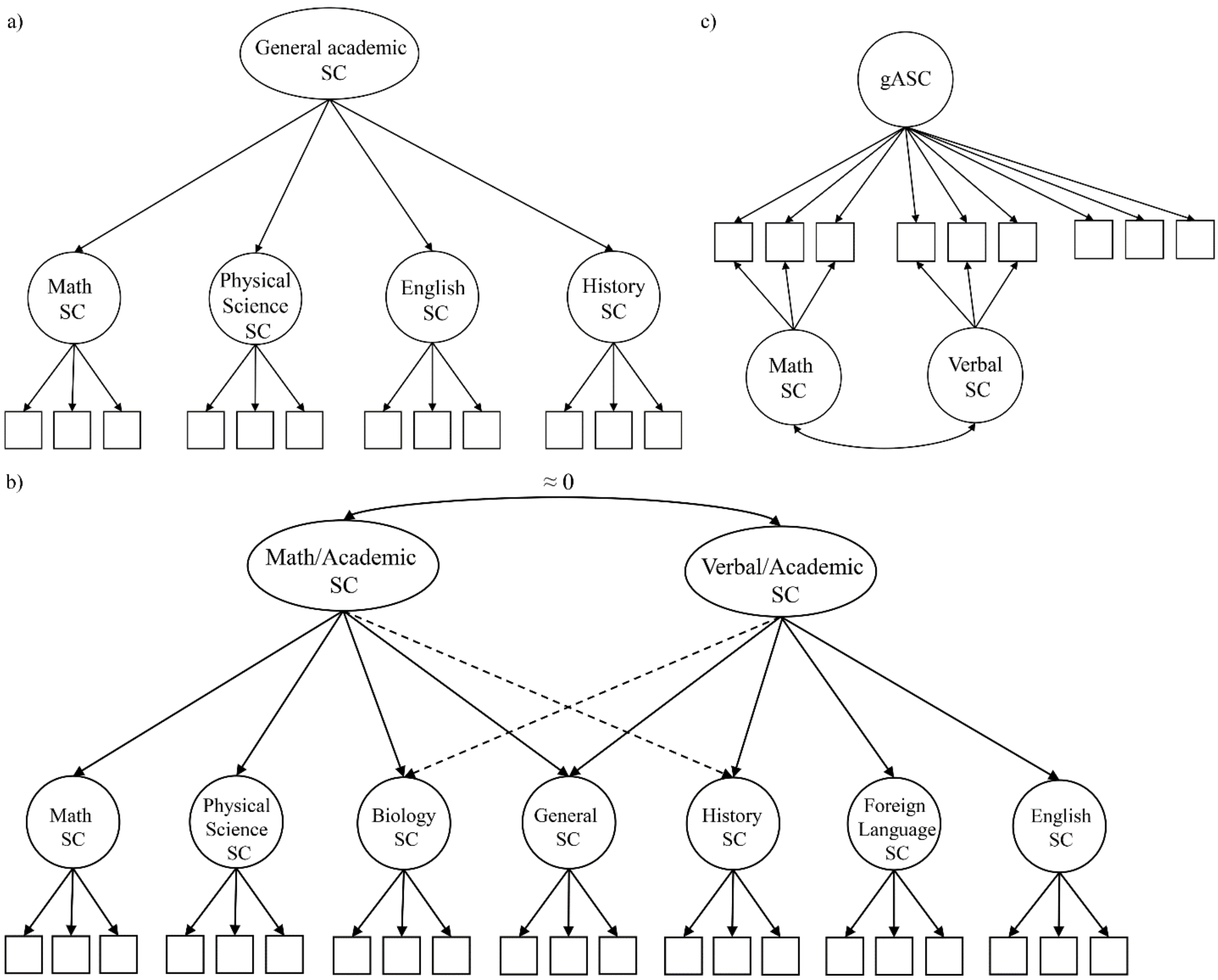


Figure 6. Structural conceptions of academic self-concept (ASC). (a) The academic section of the original model by Shavelson et al. (1976) (b) an elaboration of the Marsh/Shavelson model, (c) an elaboration of the nested Marsh/Shavelson model (Brunner et al., 2009) (a) and (b) based on Marsh et al. (Marsh et al., 1988) gASC = general ASC; SC = self-concept. To ensure clarity of presentation, residual terms of the manifest variables were omitted.

4.2.1 Different approaches to model the structure of academic self-concept

Shavelson, Hubner, and Stanton (1976) posited self-concept as a multidimensional (multifaceted) and hierarchical construct, with substantially correlated different facets of ASC forming a single higher-order factor representing general ASC (Shavelson model; see Figure 1). The multidimensionality of ASC can be depicted within a first-order correlated factor model in which the different subject-specific ASC are distinct but correlated factors. However, subsequent research showed that within the first-order correlated factor model of ASC, math ASC and native language (verbal) ASC were nearly uncorrelated. This finding was incompatible with the assumption that a single higher-order factor accounts for the correlation between the first-order factors. Thus, the Shavelson model was revised, leading to the Marsh/Shavelson model (Marsh, 1990e; Figure 6b). This model comprises two higher-order ASC factors (math/academic and verbal/academic) instead of one general higher-order factor. Moreover, general ASC is subordinate and influenced simultaneously by the two higher-order ASC factors. Several studies identifying positive correlations of general ASC with both native language and math ASC (e.g., Marsh, 1986; Marsh et al., 1988) led Brunner et al. (Brunner et al., 2008; Brunner et al., 2009; Brunner et al., 2010) to propose a new structural model of ASC, which they called the nested Marsh/Shavelson model (NMS model; Figure 6c). The NMS model combined the assumption of a general ASC at the apex and the multidimensional nature of ASC. In this model, general ASC (gASC) exerts a direct influence on all indicators of ASC and its facets—that is, general as well as all subject-specific measures. Hence, gASC is the most general factor in the NMS model, which is in line with its original conception in the Shavelson model (Shavelson et al., 1976). The subject-specific ASC factors (e.g., subject-specific math ASC or subject-specific verbal ASC) are nested under gASC and represent the multidimensional nature of ASC. In the NMS model, gASC is uncorrelated with the subject-specific ASCs, whereas the subject-specific ASCs are allowed to correlate. Thus, both gASC and subject-specific ASC explain incremental variance in subject-specific ASC measures such that their variability is decomposed into a part shared with gASC and a part unique to each subject-specific ASC. Thus, subject-specific ASC factors are residualized while the variance attributable to a gASC factor is controlled for. The correlation between the different subject-specific ASCs indicates how students contrast their strengths and weaknesses in various subjects against each other, thereby providing an insight into their profile formation. Given that gASC is controlled for in measures of subject-specific ASC, the NMS model allows researchers to investigate the profile formation of subject-specific ASC independent of the general ASC level. For example, previous research on the NMS model consistently found negative correlations between subject-specific ASC in math and verbal domains, which demonstrates that people think of themselves as either “math” or “verbal” persons (e.g., Brunner et al., 2008; Brunner et al., 2009; Brunner et al., 2010; see also Marsh & Hau, 2004). Notably, the NMS model usually comprises indicators for general ASC but does not include a specific (nested) factor for these indicators. Unlike the subject-specific indicators, the general indicators therefore load on one factor only (i.e., gASC). Hence, the model specification of the NMS model has the advantage that gASC, as the general factor, is psychometrically well defined by these indicators, which implies that the substantive interpretation of gASC as well as the correlational patterns between subject-specific ASC factors do not depend on the number of academic subjects investigated (see Eid, Lischetzke, Nussbeck, & Trierweiler, 2003; Nussbeck, Eid, Geiser, Courvoisier, & Lischetzke, 2009). Taken together, compared to a first-order correlated factor model depicting the multidimensional nature of ASCs only, the NMS model accounts for the multidimensional nature and the hierarchical structure of ASC. In a representative sample of Luxembourgish 8th graders, the NMS model described the structure of ASC well, and better than the hierarchical Shavelson model with one second-order factor of general ASC or the Marsh/Shavelson model (Brunner et al., 2010). In a cross-cultural sample of 15-year-old students in 26 countries, a restricted version of the NMS model using measures of general ASC, math ASC, and verbal ASC accounted well for the structure of ASCs (Brunner et al., 2009). Yet up to now, the NMS model has never been investigated in elementary school students.

4.2.2 Profile formation of academic self-concept in elementary school students within the framework of different structural models

Young children usually report unrealistically high levels of ASC, which has been explained by their lower abstract thinking skills, compared to older children (Harter, 1999). With cognitive development and environmental changes, reports of ASC become more realistic such that the mean ASC level decreases with age, and ASC correlates more strongly with external indicators of competence (Marsh, 1990b; Marsh, Barnes, Cairns, & Tidman, 1984; Marsh, Craven, & Debus, 1991). In particular, this development relates to two abilities: the ability to understand, interpret, and integrate ability-related experiences (such as feedback from significant others, e.g., teachers, parents, peers), and the ability to integrate information from social comparisons into one's ASC (as well as the increasing use of such information with age) (Nicholls, 1979; cf., Wigfield & Eccles, 2002). Furthermore, environmental changes (e.g., increasing competition) and more salient feedback practices (e.g., grading) contribute to the development of a more realistic ASC during elementary school (Wigfield & Eccles, 2002). In addition, an individual's grasp of self-concept changes from concrete descriptions of behavior in early childhood to trait-like psychological constructs in middle childhood, and to more abstract constructs during adolescence (Harter, 1999). From middle to late childhood, children are able to integrate opposing but coexisting concepts into their self-concept (e.g., to be dumb and smart at the same time). During late childhood, comparisons shift from those between different areas (e.g., students, friend) to comparisons between domains within the same area (e.g., in school: comparison of one's own ability in different school subjects), and to comparisons within a single domain (e.g., reading and writing in one's native language). Therefore, with increasing age, dimensional comparison processes become more differentiated and more relevant for children's views of themselves (Harter, 1999). Studies investigating the differentiation of self-concept showed that from age 4 to 5 onwards, children were able to discriminate between the different facets of ASC (e.g., reading ASC and math ASC)

(Marsh et al., 2002; Marsh, Craven, & Debus, 1998). Summarizing the findings from several studies using the Self Description Questionnaire (SDQ-I; Marsh, 1988), Marsh (1989) concluded that correlations between math ASC and reading ASC were substantial in grades 2 and 3 (*r*s = .46/.47) but low from grade 4 onwards. Besides comparing the relationship between the different ASC, Marsh (1989)found evidence for the differentiation of the different facets of ASC from grades 2 to 4, but not beyond. Furthermore, previous research indicates that whereas the correlation between math and reading ASCs decreased over time, the correlations between general ASC and math or reading ASC were moderate in size and remained relatively stable across different ages (e.g., Marsh et al., 1998; Marsh & Hocevar, 1985). One possible explanation for these complex age-related changes of correlational patterns among the different facets ASCs is Marsh and Ayotte's differential distinctiveness hypothesis (Marsh & Ayotte, 2003). It predicts an increasing differentiation of disparate areas of ASC but also an increasing integration of closely linked areas of ASC with age and cognitive development (Marsh & Ayotte, 2003, p. 689). Marsh and Ayotte summarized that "self-concept factors that were more closely associated in the self-concept hierarchy (e.g., those associated with the same higher-order factor) were predicted to show the least (or no) decline in the sizes of the correlations as age increased" (Marsh & Ayotte, 2003, p. 691). The differential distinctiveness hypothesis is provided by cross-sectional data spanning grades 2 to 6 (Marsh & Ayotte, 2003). Specifically, the correlation between reading and math ASC, which pertains to different higher-order factors in the Marsh/Shavelson model of ASC, was lower in older than in younger children, whereas their correlations with general ASC remained relatively stable. However, Marsh and Ayotte (2003) made no assumptions about changes in the correlation among the facets of skill-specific ASC of the verbal ASC, such as reading ASC and writing ASC. A few studies investigated the hierarchy within and multidimensionality of verbal ASC in samples comprising students from secondary school and beyond (e.g., Arens & Jansen, 2016; Arens, Yeung, & Hasselhorn, 2013; Lau, Yeung, Jin, & Low, 1999; Seeshing Yeung et al., 2000). They found evidence for a hierarchical and multidimensional verbal ASC, with a general verbal ASC as a higher-order factor as well as skill-specific verbal ASCs (i.e., reading, writing). For eighth- and ninth-grade students as well as for university students, correlations between different skill-specific verbal ASC were high (*r*s ranging between .53 and .87). Following Marsh and Ayotte's train of thought that the relationship between self-concepts belonging to the same higher-order factor would strengthen with age (Marsh & Ayotte, 2003), the correlation of reading ASC and writing ASC, as skill-specific self-concepts of the verbal domain, should at least be stable or even increase with age. To date, this has not been explicitly tested. In general, reading ASC has been examined most frequently as a single indicator of verbal ASC. In contrast, writing ASC has received comparably little attention in research on the structure of ASC, although research has indicated that reading ASC does not cover verbal ASC sufficiently (Arens, Yeung, & Hasselhorn, 2013). In elementary school education, both reading and writing are core skills that students have to master (Ehm et al., 2014). Previous research has found reading ASC and writing ASC to be distinct but correlated factors within verbal ASC. For instance, Ehm et al. (2014) examining third graders, showed that a first-order correlated factor model of ASC with writing ASC, reading ASC, and math ASC provided a better fit to the data than a model with one combined verbal ASC factor. In this study, the correlation between reading ASC and writing ASC was substantial (r = .66). Poloczek, Karst, Praetorius, and Lipowsky (2011) found a similar correlation between reading ASC and writing ASC (r = .58) for first grade elementary students. Taken together, these findings indicate positive correlations of similar size between reading ASC and writing ASC in samples of children, adolescents, and adults. It should be noted that, for the most part, previous research investigating the differentiation of self-concept has compared the sizes of correlations between different self-concept facets using factor scores from first-order confirmatory factor analysis, factor scores from exploratory factor analysis, or manifest scale scores. However, in these studies, the overall level of subject-specific ASC confounds variability unique to the subject-specific ASC and variability due to general ASC. When investigating students' ASC profile formation as one key aspect of ASC differentiation, it is advantageous to use factor scores derived from the NMS model, which do not suffer from this confound because the general level of ASC is controlled for in subject-specific measures (see Figure 6c). Thus, differentiation in students' profiles of subject-specific ASC is represented more clearly in the NMS model than in the first-order correlated factor model, where the two sources of variance are not separated so that subject-specific ASC factors still contain both. Therefore, correlational patterns among the different subject-specific ASCs are more ambiguous to interpret because changes in these correlations may be attributable to (a) changes in students' ASC profile or (b) changes in the relation of subject-specific measures to gASC, that is, the amount of variability in subject-specific ASCs explained by general ASC. To conclude, no study has yet investigated profile formation of ASC within the NMS model of ASC based on the assumptions of the differential distinctiveness hypotheses considering the different subject-specific ASCs as well as skill-specific ASC in elementary school students. Available research findings suggest that with increasing age, disparate areas of ASC (e.g., reading ASC and math ASC) correlate less, whereas correlations between closely linked areas of ASC (e.g., reading ASC and writing ASC) remain stable or even increase.

4.2.3 The internal/external frame of reference model in elementary school students and the role of achievement differences across domains in the I/E model

The internal/external frame of reference model (I/E model; Marsh, 1986) states that external and internal (dimensional) comparison processes play an important role in the formation of ASC. First, students make use of external comparisons: They compare their performance in a particular subject with the performance of other students in the same subject and with further external standards of actual achievement level. Second, students make use of dimensional comparisons, which occur when students contrast their performance in one particular school subject with their performance in other subjects. External comparisons lead to positive correlations between achievement and the corresponding ASC within a domain, whereas dimensional comparisons can either reduce correlations between subject-specific ASCs through contrast effects or increase correlations through assimilation effects. Typically, the assumptions of the I/E model are tested via path models. In a path model, the effects of external comparisons are evidenced by positive paths from achievement to the student's ASC in the corresponding subject. For dimensional comparisons, assimilation effects are evidenced by positive paths from achievement in one subject to the ASC in the non-corresponding subject. Contrast effects, on the other hand, are evidenced by negative paths between achievement and ASC of non-corresponding subjects. Summarizing previous research on dimensional comparison processes, Möller and Marsh's dimensional comparison theory (see Möller & Marsh, 2013) posits that contrast effects are more likely when subjects are dissimilar. Based on the similarity/dissimilarity of subjects, Marsh et al. (2014) arranged the subject-specific ASCs (e.g., math ASC, biology ASC, or foreign language ASC) along a continuum that places ASCs of similar subjects close to each other (see Figure 6b). Math and verbal (reading/writing) ASCs as the most dissimilar subjects are located at the opposite ends of this continuum. Contrast effects are more likely between "distant" subjects, whereas assimilation effects are more likely between "close" subjects or subjects pertaining to the same domain (e.g., reading/writing). Thus, comparisons of achievement in math and native language (e.g., reading/writing) should result in the largest contrast effects. Most studies have investigated the I/E model with the subjects math and native language. For this case of subjects from dissimilar domains, cross paths leading from math achievement to verbal ASC and vice versa are predicted to be negative (i.e., contrast effects). A large number of studies on math and verbal subjects support the I/E model. However, most of those studies have focused on secondary school students. In their meta-analysis of I/E model findings, Möller et al. (2009) noted that elementary school students were examined in only 3 out of 69 samples. Moreover, none of these 3 studies covered both reading and writing ASC, although previous research showed the two to be separate factors of verbal ASC (see above). Hence, in more recent studies, the I/E model was extended. For instance, Ehm et al. (2014) investigated the I/E model with reading ASC, writing ASC, and math ASC in elementary school grades 1, 2, and 3. They found evidence for dimensional comparisons only in 3rd grade, and this for reading and math only. Poloczek et al. (2011) examined the I/E model with math and reading in first-graders and also found no effects of dimensional comparisons for this grade level. These findings indicate that effects of dimensional comparisons are likely to occur from 3rd grade onwards. Most I/E studies have either neglected the general ASC factor in favor of domain-specific ASC factors or included a general ASC and a general achievement indicator as correlated factors to control for their influence while studying the I/E model (e.g., Marsh, 1986). An exception is a study by Brunner and colleagues (Brunner et al., 2008; Brunner et al., 2010) which incorporated general ASC using the NMS model and contrasted the results with the common test of the I/E model. Across both ASC models, the effects of external and internal comparisons assumed by the I/E model could be supported. However, no study has yet modeled the assumptions of the I/E model by applying the NMS model to elementary school students. Marsh, p.110 (1990c) framed the predictions of the I/E model as general effects because lower self-perceived skills in one's native language will lead to higher levels of math ASC for students across the entire math skill spectrum and vice versa. However, these effects may be reinforced by an interaction of achievement in both subjects. That is, the effect of internal comparisons may increase when achievement differences in contrasting subjects become larger and thus more salient. For example, Rost et al. (2004; 2005) using a quasi-experimental approach, found that effects of internal comparisons were stronger when achievement levels diverged in the different subjects. In a sample of German students (7^{th} /8^{th} grade), Rost et al. categorized students into different groups: students with identical grades in all subjects (Group 1), students with a one-grade difference between subjects (Group 2), and students with a greater-than-one grade difference between subjects (Group 3). In Group 1, the correlation between the different subject-specific ASCs was substantial and positive, whereas it was lower in Group 2, and lowest in Group 3. Whether this moderating effect of achievement differences for dimensional comparisons generalizes to students in elementary school remains an open question.

4.3 Research aims of the present study

In the present study, we capitalized on data from a quasi-representative cross-sectional sample of German elementary school students in grades 1 to 4. We pursued three major research goals: First, we tested the suitability of the NMS model of ASC in elementary school students in comparison to other structural models. For this age group, there is evidence for a general ASC, for different subject-specific ASCs (Marsh, 1986), and for distinct writing and reading factors of verbal ASC (Arens & Jansen, 2016; Ehm et al., 2014). Therefore, an NMS model of ASC (Brunner et al., 2009) with math ASC, writing ASC, and reading ASC as three correlated specific factors nested under a general ASC (gASC; Model 1) should provide a good fit to the data of elementary school students. Furthermore, a first-order correlated factor model with general ASC and math ASC, writing ASC, and reading ASC should fit the data well, too (Model 3). For the sake of completeness, we compared these two models to the following models: an NMS model with reading and writing collapsed into a single verbal ASC factor (Model 2); a first-order correlated factor model with general ASC, math ASC, and one verbal ASC factor (Model 4); a g-factor model in which all indicators (i.e., items) of general ASC, math ASC, reading ASC, and writing ASC were explained by one first-order factor (Model 5). All models tested in this study are shown in Figure 7. Second, we investigated how ASC profiles form in elementary school students. We assumed profile formation to take place early in elementary school (from middle to late childhood; e.g., Harter, 1999; Marsh et al., 1991). In line with the differential distinctiveness hypothesis (Marsh & Ayotte, 2003), we expected to find (a) an increasing integration of closely related areas of self-concept (e.g., writing/reading ASC), resulting in stable or increasing positive correlations with increasing grade level. Further, we expected to find (b) a differentiation of diverging areas of self-concept (e.g., math/writing ASC and math/reading ASC), resulting in decreasing positive correlations moving to zero or becoming negative and increasing in size with increasing grade level. Previous studies mainly used a first-order correlated factor model of ASC to investigate this research question. To align our results with those from previous research, we investigated profile formation of ASC both within the first-order correlated factor model of ASC and within the NMS model (see Figure 7: Model 1 and Model 3). In the NMS model, variability in subject-specific ASC measures is residualized by controlling for the variance attributable to gASC. Accordingly, correlations among subject-specific ASCs ("specific factors") directly depict the profile formation of subject-specific ASCs independent of the general level of the ASC profile. Therefore, we expected the change in correlations between the domain-specific ASCs to become more apparent in the NMS model than in the first-order correlated factor model. Third, to learn more about the processes underlying profile formation of ASC, we tested the I/E model within the NMS model and the first-order correlated factor model using math and German grades as achievement indicators. Figure 8a shows the I/E model within the NMS model; Figure 8b illustrates the I/E model within the first-order correlated factor model. In grades 3 and 4, we extended the I/E model by including an interaction term (math grade × German grade) as a predictor of math ASC, reading ASC, and writing ASC to test our assumption that effects of dimensional comparisons become stronger when achievements in cross-domains differ (Rost et al., 2004; Rost et al., 2005)

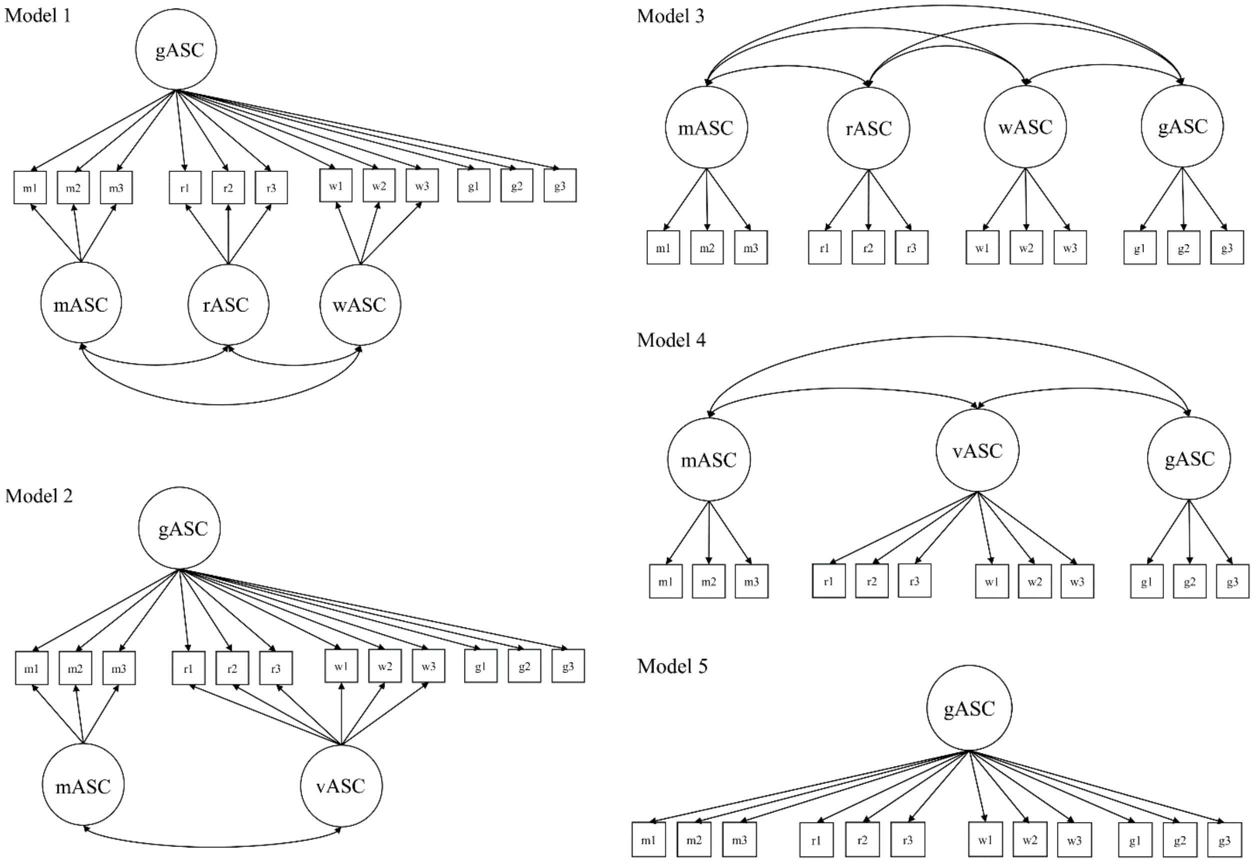
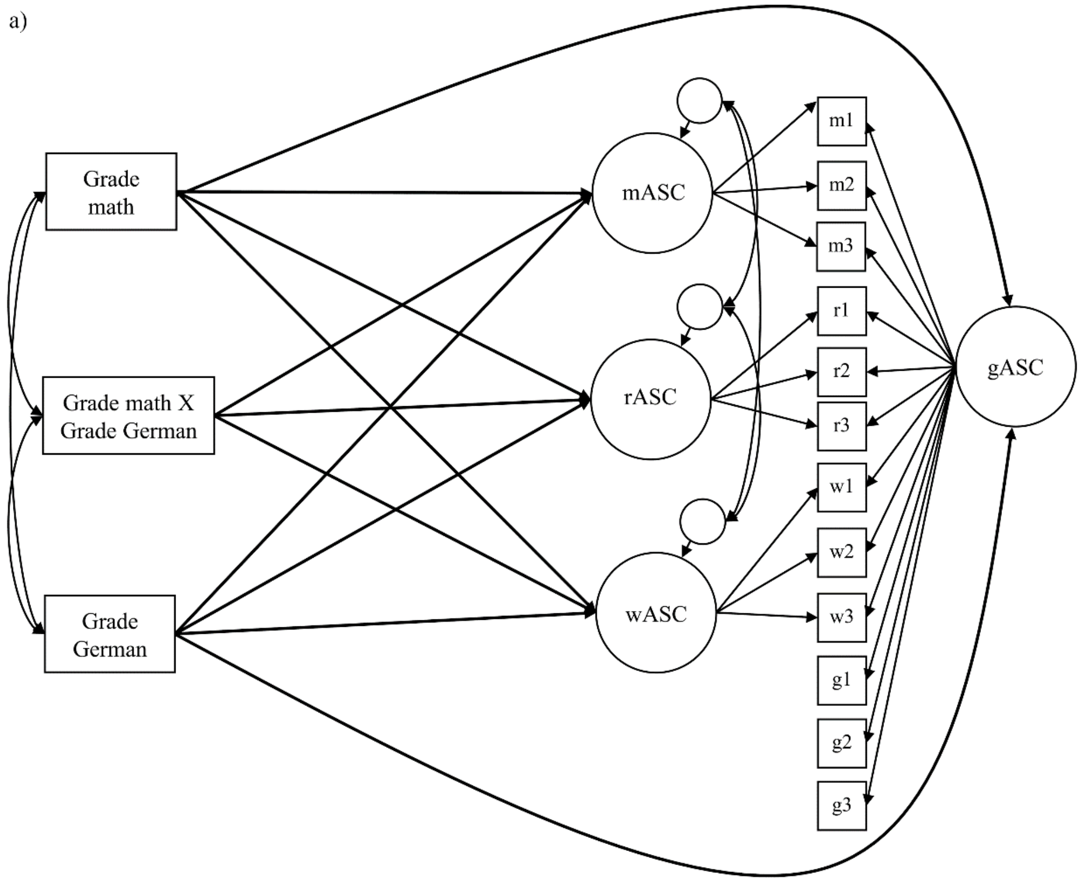


Figure 7. Model 1: NMS_3sf = Nested Marsh/Shavelson model of academic self-concept (ASC) with math ASC, writing ASC, and reading ASC as correlated specific factors. Model 2: Nested Marsh/Shavelson model of ASC with math ASC and verbal ASC as correlated specific factors. Model 3: FOCF_4f = First-order correlated factor model with general ASC, math ASC, reading ASC, and writing ASC. Model 4: FOCF_3f = First-order correlated factor model with general ASC, math ASC, and verbal ASC. Model 5 = *g*-factor model in which all ASC items load on one factor. gASC = general ASC, mASC = math ASC, wASC = writing ASC, rASC = reading ASC, vASC = verbal ASC. For the sake of clarity, residual terms of the manifest variables were omitted.

We expected a comparable size of the effect in grade 3 and grade 4 because the moderating effect of achievement differences should depend on children's ability to evaluate the size of the difference between their grades in math and the native language. Research indicates that children are already able to do so at the beginning of elementary school (from age 6 onwards; e.g., Krajewski, Nieding, & Schneider, 2008). To the best of our knowledge, our study is the first to extend the I/E model by an interaction term capturing the impact of grade differences in elementary school students.



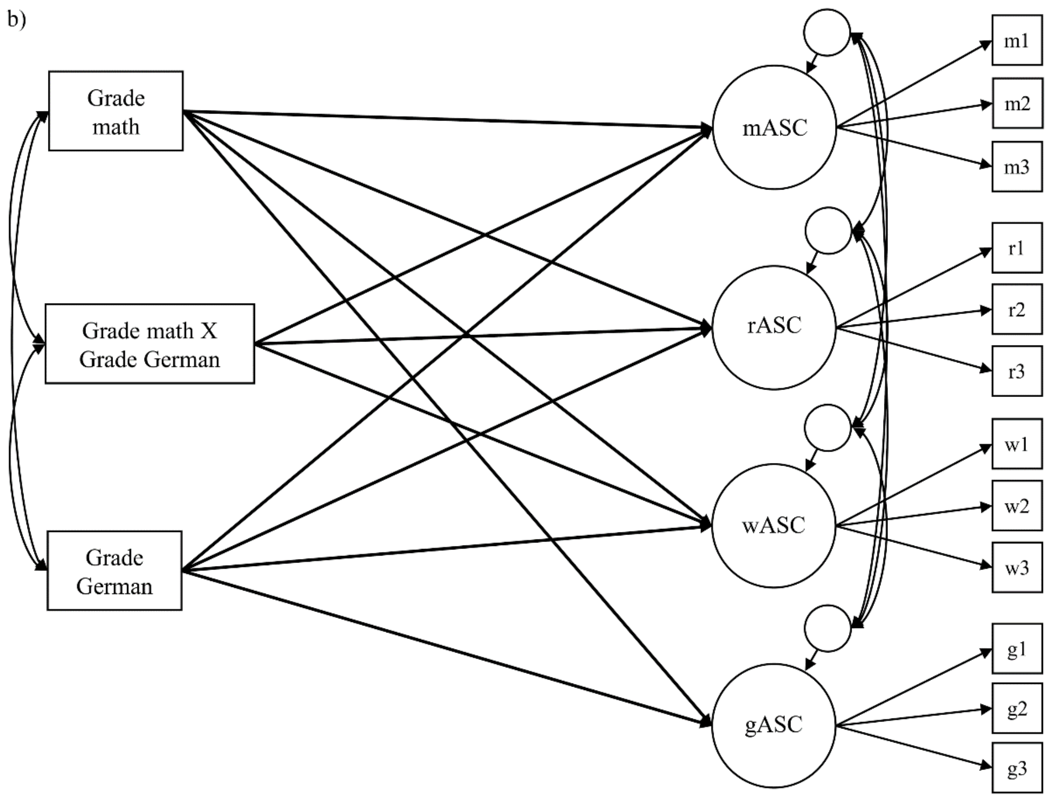


Figure 8. Extended I/E-Model. (a) I/E-NMS model. gASC = general ASC, mASC = math ASC as specific factor, wASC = writing ASC as specific factor, rASC = reading ASC as specific factor.

(b) I/E-FO model. gASC = general ASC, mASC = math ASC, wASC = writing ASC, rASC = reading ASC as first-order factor. For the sake of clarity, residual terms of the manifest variables were omitted.

4.4 Method

4.4.1 Sample and procedure

The sample is composed of a quasi-representative cross-sectional sample of German elementary school students (the norming sample of the intelligence test "THINK1–4"; Baudson, Wollschläger, & Preckel, 2016). The data used in this study were collected in five German federal states (Hesse, Lower Saxony, Mecklenburg-West Pomerania, North Rhine-Westphalia, and Rhineland-Palatinate) between September 2012 and January 2014. Thus, we took care to cover a broad range of federal states because of the north south as well as east-west disparity in Germany; also, we ensured that rural and urban areas were adequately covered. Permissions for the different federal states were obtained mainly through contacts with the network of the editors of the series "Hogrefe Schultests". The total sample comprised N = 3,779 students (50.2% female). The grade 1 sample consisted of 427 students (47.3% female) from 34 classes, the grade 2 sample consisted of 718 students (50.8% female) from 53 classes, the grade 3 sample consisted of 1,288 students (52.6%) from 82 classes, and the grade 4 sample consisted of 1,346 students (49.5% female) from 90 classes. Students' mean age was 6.43 years (SD = .51) in grade 1, 7.63 years (SD = .57) in grade 2, 8.67 years (SD = .60) in grade 3, and 9.66 years (SD = .59) in grade 4. The larger samples in grades 3 and 4 were due to oversampling by design (the project THINK further addressed the transition to secondary school). In all grade levels, ASCs were assessed by self-report. Student questionnaires were administered in class. The study was approved by the responsible authorities of the federal states of Germany. This permission required the approval of the data protection commissioners of each respective federal state. In Hesse, the study was approved on March 7, 2013 by the Hessian Ministry of Education, No.: 660.003.000–00544. In Lower Saxony, the study was approved on August 16, 2012 by the Lower Saxony State Board of Education, No.: BS 1 R.21-81402/1-26/12. In Mecklenburg-West Pomerania, the study was approved in August 2012 by the Ministry of Education, Science and Culture of Mecklenburg-West Pomerania, No.: VII 201 g. In Rhineland-Palatinate, the study was approved on May 4, 2012 by the Supervision and Service Administration of Rhineland-Palatinate, No.: 51 111-32/70-12. In North Rhine-Westphalia, permissions were given orally by the school principals who have the decision-making power in this federal state of Germany. Student participation was anonymous and voluntary. Prior written parental consent was obtained for all participating children. Anonymity was ensured by predefined codes that were assigned by teachers, according to coding instructions. Only teachers knew which code referred to which child. The same code was assigned to each set of questionnaires completed by one family (i.e., child-parent dyads). On average, 68% of all students in each class participated in the study. The overall consent rate was 79%. Dropout was due to missing parental approval or sickness on the day of the examination. Data collection took place in all four quarters of the school year. Within each school, data were collected in the same quarter, but quarters varied between schools. There was no significant correlation between the quarters and the ASC scale means, with the exception of writing ASC in grade 3 (r = -.13, p = .016).

4.4.2 Variables and measures

Academic self-concepts. To assess general ASC and writing, reading, and math ASC, we applied 3-item scales for each conception from the short version of the FEESS 1–2 & 3–4 (Rauer & Schuck, 2003, 2004). The items were: “I am good at reading/writing/math/in school” (r1, w1, m1, g1), “I do well in reading/writing/math/school” (r2, w2, m2, g2), and “In reading/ writing/math/school, I do most things well” (r3, w3, m3, g3). To facilitate responses for the first and second graders in the sample and to make ratings comparable across grade levels (the original FEESS uses a dichotomous scale in grades 1 and 2 and a four-point rating scale in grades 3 and

4), a three-point rating scale was employed with frowning, neutral, and smiling faces to indicate disagreement, indifference, and agreement. A smiley scale is easy to understand and suitable for the assessment with elementary school children (see Davies & Brember, 1994). For all students, a sample task was used to explain the smiley scale. Taking into account that not all students were yet able to read fluently in grades 1 and 2, all items were read aloud to these students.

School grades. For participants in grades 3 and 4, parents reported their children's teacher-assigned school grades at the time of data collection (i.e., from their most recent report card), while no grades were reported for 1st and 2nd graders because children are not graded in Germany until grade 3. All parent reports were anonymous due to the teacher-assigned codes; parent questionnaires were returned in sealed envelopes, and teachers did not have access to the returned data. The instructions requested parents to provide valid information. Thus, we trusted the validity of the parent-reported grades. For the present analyses, the usual German grading scale (from 1 = very good to 6 = insufficient) was inverted so that higher numbers indicate higher achievement.

4.4.3 Data analyses

First, we conducted confirmatory factor analyses (CFA) to investigate the different structural models of ASC. All model parameters were analyzed with the statistical software Mplus 7.31 (Muthén & Muthén, 1998-2017). As is usually the case in educational research, students were nested within classes, which may result in distorted significance tests. Hence, we used the COMPLEX option in Mplus to obtain correct standard errors of parameter estimates and test statistics. To account for the ordered categorical nature of the data (three-point rating scales), we conducted CFAs for ordered categorical data using the mean- and variance-adjusted weighted least square estimator (WLSMV) (Muthén, Du Toit, & Spisic, 1997) with theta parameterization. To evaluate model fit of the CFAs, we used the chi-square ($\chi 2$) goodness-of-fit statistic. Because this statistic is sensitive to sample size, we also used the following recommended descriptive measures of model fit (Hu

& Bentler, 1999): (1) the root mean square error of approximation (RMSEA), which should be below .06, (2) the comparative fit index (CFI), which should exceed .95, and (3) the weighted root mean residual (WRMR), which should be below 1 (Yu, 2002). We compared the NMS model with a gASC and reading ASC, writing ASC, and math ASC as correlated specific ASC factors nested under gASC (Model 1: NMS_3sf; sf for "specific factors") against alternative structural models: an NMS model with math ASC and one comprehensive verbal ASC as specific ASC factors (Model 2: NMS_2sf); a first-order correlated factor model with general ASC, math ASC, reading ASC, and writing ASC as distinct but correlated factors (Model 3: FOCF_4f; f for "factors"), a first-order correlated factor model with general ASC, math ASC, and only one verbal ASC factor (Model 4: FOCF_3f), and a g-factor model in which all items were explained by a single first-order factor (Model 5: g-factor) (see Figure 7). Latent variables in all structural models of ASC were identified by fixing the first item loading of each factor to 1. Further, covariances between general ASC and the two or three subject-specific ASC factors were constrained to 0 in the NMS_3sf model and the NMS_2sf model. This implies that subject-specific ASC factors are truly residualized while the variance attributed to gASC is controlled for (Brunner et al., 2010). All other model parameters were freely estimated. The models were tested separately for grades 1 to 4. Second, we tested the NMS_3sf and FOCF_4f models (i.e., Models 1 and 3) for scalar measurement invariance across grade levels through multigroup comparison. This is the best-practice approach to group comparisons regarding the structural models of ASC and a prerequisite for further invariance testing of correlations and factor means of ASCs (Sass, 2011). The measurement invariance testing procedure consisted of comparing nested models differing in number of parameters, which were restricted to be invariant across groups (see, e.g., Sass, 2011). In the configural model, the same structural model was estimated simultaneously for each group, but the estimated parameters (factor loadings and thresholds) were allowed to vary between groups. In the scalar model, factor loadings for identical items and their corresponding thresholds were held equal across grades (Sass, 2011). To assess scalar invariance, we used the $\chi2$ difference test between the configural and the scalar model using the DIFFTEST option in Mplus. It should be noted that with ordered categorical indicators and indicators loading on two factors, it is not possible to test the invariance of the factor loadings (i.e., metric measurement invariance) separately from scalar measurement invariance (see, e.g., Millsap & Yun-Tein, 2004). Third, under the assumption of scalar measurement invariance, we tested the hypothesis that ASC profiles are more distinct in higher grades by comparing all pairwise correlations among math ASC, reading ASC, and writing ASC between the 4 grade levels using the Wald $\chi2$ test. To interpret the magnitude of correlations, we used the classification proposed by Cohen (1992), with *r*s of .10, .30, and .50 indicating small, medium, and large effects, respectively. Fourth, to investigate the assumptions of the extended I/E model, we conducted separate path analyses in grades 3 and 4. Math and German grades were included into the model as manifest variables. To avoid multicollinearity, we standardized grades beforehand within each school year and computed the interaction term between the grades in math and German. To obtain fully standardized path coefficients in the I/E model, we defined the metric in the NMS_3sf and FOCF_4f models by fixing the variance of the latent factors to 1. Additionally, to account for general achievement (equivalently to general ASC), we added paths leading from grades in German and math to the general ASC factor. Figure 8a shows the extended I/E model with the NMS_3sf model of ASC (I/E-NMS model). Figure 8b shows the extended I/E model with the FOCF_4f model of ASC (I/E-FO model).

4.5 Results

4.5.1 Descriptive statistics

Table 3 shows the means, standard deviations, and manifest correlations of all measures by grade level. In line with previous findings, descriptive statistics indicated lower mean levels of reading

ASC, writing ASC, math ASC, and general ASC with increasing grade level (except for reading ASC between grades 1 and 2). Furthermore, at a descriptive level, manifest correlations between writing/reading ASC and math ASC decreased with increasing grade level; the correlation between writing and reading ASC was lower in grade 2 than in grade 1, and higher in grade 4 than in grade 3. Scale reliabilities in terms of Cronbach's alpha (α) were acceptable to good, ranging from α = .68 to .87 across grade levels.

Table 3. Means, standard deviations, and manifest correlations of measures by grade level

Grade 1 (above diagonal) and Grade 2 (below diagonal)								
	ASC general	ASC math	ASC writing	ASC reading	*n*	*M*	*SD*	α
ASC general	--	.52**	.42*	.41***	422	1.73	.41	.68
ASC math	.51**	--	.44**	.38***	365	1.75	.44	.83
ASC writing	.44**	.23**	--	.53**	367	1.75	.42	.77
ASC reading	.40**	.22**	.52**	--	363	1.73	.47	.81
n	716	356	356	370				
M	1.69	1.69	1.69	1.76				
SD	.40	.46	.46	.42				
α	.74	.83	.81	.77				

Grade 3 (above diagonal) and Grade 4 (below diagonal)										
	ASC general	ASC math	ASC writing	ASC reading	Grade math	Grade German	*n*	*M*	*SD*	α
ASC general	--	.52**	.44*	.42***	.35**	.31**	726	1.60	.46	.74
ASC math	.53**	--	.17**	.16***	.37**	.04	334	1.60	.51	.86

ASC writing	.43**	.15**	--	.35**	.14*	.29**	333	1.54	.51	.82
ASC reading	.40**	.13**	.40**	--	.20**	.31**	338	1.70	.44	.82
Grade math	.46**	.54**	.10**	.14**	--	.65**	1,174	5.0	.76	
Grade German	.49**	.23**	.38**	.38**	.67**	--	1,182	4.9	.74	
n	1,255	854	862	370	1,173	1,169				
M	1.56	1.58	1.46	1.67	4.82	4.69				
SD	.44	.49	.54	.44	.88	.85				
α	.74	.86	.87	.87						

Note. Descriptive statistics were computed using IBM SPSS Statistics version 22.0 (SPSS Inc., USA). Scale reliability was calculated using Cronbach's α. † $p < .10$ * $p < .05$ ** $p < .01$ *** $p < .001$.

4.5.2 Fit of the different structural models of ASC in grades 1 to 4

Table 4 shows the model fit results of the five different structural models of ASC by grade level. As expected, the NMS_3sf and the FOCF_4f models fitted the data well. In both models, latent constructs were measured adequately by their respective indicators at all grade levels (standardized factor loadings: NMS_3sf > .47, all *p*s < .001; FOCF_4f model > .68, all *p*s < .001; see Appendix B, Table B 1). The models assuming only one verbal ASC factor (NMS_2sf and FOCF_3f) showed worse fit than the models distinguishing reading ASC and writing ASC. The g-factor model of ASC showed insufficient fit. With reference to McDonald (McDonald, 1999) (see also Brunner, Nagy, & Wilhelm, 2012), we used omega (ω) as the reliability coefficient for NMS_3sf and FOCF_4f, the two structural models of ASC favored by our results. ω can be interpreted as a measure of internal consistency within a latent variable approach. Scale reliability was excellent, ranging from .81 to .96 for both models (see Appendix B, Table B 2).

Table 4. Model fit results of the five structural models (M1: NMS_3sf; M2: NMS_2sf; M3: FOCF_4f; M4: FOCF_3f; M5: g-factor) of academic self-concept by grade level

Model	N	χ^2	df	p	CFI	RMSEA [90% CI]	WRMR
Grade 1							
M1: NMS_3sf	427	45.811	42	.317	.999	.015 [.000;.037]	.477
M2: NMS_2sf		100.129	44	<.001	.980	.055 [.041;.069]	.860
M3: FOCF_4f		73.069	48	.011	.991	.035 [.017;.050]	.701
M4: FOCF_3f		133.571	51	<.001	.970	.062 [.049;.074]	1.113
M5: *g*-factor		235.277	54	<.001	.935	.089 [.077;.100]	1.791
Grade 2							
M1: NMS_3sf	718	45.598	42	.325	.999	.011 [.000;.028]	.491
M2: NMS_2sf		136.243	44	<.001	.975	.054 [.044;.064]	1.109
M3: FOCF_4f		78.234	48	.004	.992	.030 [.017;.041]	.776
M4: FOCF_3f		148.701	51	<.001	.973	.052 [.042;.061]	1.330
M5: *g*-factor		410.464	54	<.001	.902	.096 [.087;.105]	2.745
Grade 3							
M1: NMS_3sf	726	58.999	42	.043	.997	.024 [.005;.037]	.590
M2: NMS_2sf		136.243	44	<.001	.975	.054 [.044;.064]	1.109
M3: FOCF_4f		112.500	48	<.001	.987	.043 [.033;.053]	1.073
M4: FOCF_3f		245.766	51	<.001	.961	.073 [.064;.082]	1.944
M5: *g*-factor		851.055	54	<.001	.839	.143 [.134;.151]	4.057
Grade 4							
M1: NMS_3sf	1259	143.424	42	<.001	.996	.044 [.036;.052]	.890
M2: NMS_2sf		965.181	44	<.001	.962	.129 [.122;.136]	3.100
M3: FOCF_4f		239.363	48	<.001	.992	.056 [.049;.063]	1.302
M4: FOCF_3f		818.168	51	<.001	.968	.109 [.103;.116]	3.238
M5: *g*-factor		2340.188	54	<.001	.905	.183 [.177;.190]	6.718

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; WRMR = weighted root mean square residual. M1: NMS_3sf = Nested Marsh/Shavelson model of ASC with math ASC, writing ASC, and reading ASC as correlated specific factors; M2: NMS_2sf = Nested Marsh/Shavelson model of ASC with math ASC and verbal ASC as correlated specific factors; M3: FOCF_4f = First-order correlated factor model with general ASC, math ASC, writing ASC, and reading ASC as factors; M4: FOCF_3f = First-order correlated factor model with general ASC, math ASC and verbal ASC as factors; M5: g-factor = First-order factor model in which all ASC items load on one single factor.

4.5.3 Measurement invariance across grade levels

Table 5 shows the results of the stepwise investigation of measurement invariance across grade levels of the NMS_3sf and the FOCF_4f models. Importantly, the non-significant $\chi2$ difference test between the configural and the scalar invariant models indicated strong factorial measurement invariance for both the NMS_3sf and the FOCF_4f model, allowing us to compare correlations of factors as well as mean differences in factor scores across grade levels (Brown, 2006). The procedure and the results of latent mean level differences tests in factor scores across grade levels are presented in Appendix B.

Table 5. Results of the measurement invariance tests for the NMS model (M1: NMS_3sf) and the first-order correlated factor model (M3: FOCF_4f), both with math ASC, writing ASC, and reading ASC as distinct factors

Model	Invariance level	χ^2	*df*	$\Delta\chi^2$ *(Δdf)*	*p*	CFI	RMSEA [90% CI]	WRMR
	configural	268.741	168		<.001	.997	.028 [.021;.034]	1.266
M1	scalar	325.141	243		<.001	.998	.021 [.014;.026]	1.556
	configural vs. scalar			87.958 (75)	.145			
	configural	463.313	192		<.001	.992	.042 [.038;.047]	1.983
M3	scalar	494.622	240		<.001	.992	.037 [.037;.032]	2.060
	configural vs. scalar			52.268 (48)	.312			

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; CI = confidence interval; WRMR = weighted root mean square residual. M1: NMS_3sf = Nested Marsh/Shavelson model of ASC with math ASC, writing ASC, and reading ASC as correlated specific factors; M3: FOCF_4f = First-order correlated factor model with general ASC, math ASC, writing ASC, and reading ASC as factors.

4.5.4 Profile formation of academic self-concept

For the NMS_3sf and the FOCF_4f models, descriptive results revealed lower correlations in higher grades among ASCs of non-corresponding domains (math and reading/writing) and within the verbal domain (reading and writing) (see Figure 8). Both sizes and signs of correlations differed between the two models. In the NMS_3sf model, reading ASC and math ASC were weakly and positively correlated in grade 1; their correlation was weak and negative in grade 2, and even more negative and of medium size in grade 3, but then remained at about the same level in grade 4. The correlation of writing ASC and math ASC showed the same pattern, except that the positive correlation was larger, that is, of medium size in grade 1. In contrast, in the FOCF_4f model, correlations between reading ASC and math ASC and between writing ASC and math ASC were positive at every grade level and lower in size in higher grades. Specifically, correlations between math ASC and writing ASC/reading ASC were high and positive in grade 1, of medium size and positive in grade 2, and weak and positive in grades 3 and 4. Regardless of the structural model investigated, correlations between writing ASC and reading ASC were always positive, and they were lower in higher grades. However, the size of the correlations was larger in the FOCF_4f model than in the NMS_3sf model. For both models, Wald $\chi2$ tests of the difference of correlations between reading ASC and math ASC were significant between grades 1 and 2, whereas the correlations between writing ASC and math ASC differed significantly between the two grades in the NMS_3sf model only. Further, in both models, significantly different correlations between reading ASC and math ASC were observed between grades 1 and 3 and between grades 1 and 4. In both models, the correlation between writing ASC and math ASC differed significantly between grades 1 and 3 and between grades 1 and 4, but not between grades 2 and 3, 2 and 4, or 3 and 4 (see Figure 9). Contrary to our expectations, which were derived from the differential distinctiveness hypothesis, both models showed significantly lower correlations between writing ASC and reading ASC in grade 3 compared to grades 2 and 1. However, correlations between writing ASC and reading ASC were slightly higher in grade 4 than in grade 3, but only on a descriptive level. In sum, in both models the pattern of results indicated that ASC profiles were formed early, that is, within the first three years of school. Profile formation for math ASC and verbal (i.e., reading) ASC occurred within the first two years of school, whereas results for the profile formation within the native language (i.e., reading ASC and writing ASC) suggested a time frame of the first three years. However, results concerning profile formation between math ASC and writing ASC differed between the two models. Results of the NMS_3sf model suggested profile formation to take place within the first two years of school, whereas results of the FOCF_4f model suggested a time frame of the first three years.

Grade	*N*	M1	M3	M1	M3	M1	M3
		mASC and rASC		mASC and wASC		rASC and wASC	
1	427	.19	.59***	.33*	.69***	.54***	.73***
2	718	-.18	.34***	-.27†	.36***	.51***	.69***
3	726	-.44***	.21***	-.43***	.20**	.19*	.50***
4	1,259	-.42***	.18**	-.43***	.19***	.26***	.51***

Figure 9. Latent correlations between academic self-concepts in math, writing, and reading by model (M1: NMS_3sf; M3: FOCF_4f) and grade level. Despite equal nomenclature, mASC = math ASC, wASC = writing ASC, and rASC = reading ASC are first-order correlated factors in the FOCF_4f model and specific factors that are residualized by gASC in the NMS_3sf model; thus, factors from both models are not directly comparable. M1: NMS_3sf = Nested Marsh/Shavelson model of ASC with math ASC, writing ASC, and reading ASC as correlated specific factors; M3: FOCF_4f = First-order correlated factor model with general ASC, math ASC, writing ASC, and reading ASC as factors.
^{t} $p < .10$ * $p < .05$ ** $p < .01$ *** $p < .001$. Gray arrows mark significant differences between correlation coefficients of different grades. How to read the table (e.g., column 4, M1, mASC and rASC): The difference in the correlation between mASC and rASC in Model 1 was significant between grades 1 and 2, 1 and 3, and 1 and 4

4.5.5 Test of the extended I/E model

We tested the extended I/E model within the NMS model (Model I/E-NMS) and within the first-order correlated factor model (Model I/E-FO) in grades 3 and 4 separately. Both models fitted the data well, though fit was slightly better for the I/E-NMS model than for the I/E-FO model (see Table 6).

Table 6. Model fit results of the extended I/E-Model within the NMS model of ASC (Model I/E-NMS) and within the first-order correlated factor model of ASC (Model I/E-FO) for grades 3 and 4

Model	*N*	χ^2	*df*	*p*	CFI	RMSEA [90%CI]	WRMR
Grade 3							
I/E-NMS	620	122.791	67	<.001	.981	.037 [.026;.047]	0.854
I/E-FO		158.859	73	<.001	.971	.044 [.034;.053]	1.093
Grade 4							
I/E-NMS	1079	294.426	67	<.001	.978	.050 [.044;.057]	1.158
I/E-FO		242.453	73	<.001	.982	.053 [.047;.059]	1.310

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation; WRMR = weighted root mean square residual.

External comparisons. In both grades and in both models, paths from school grades to the specific ASC in the corresponding domain were significant and positive, indicating significant effects of external comparisons (see Table 7).

Table 7. Standardized path coefficients for the extended I/E Model within the NMS model of ASC (I/E-NMS Model) and within the first-order correlated factor model of ASC (I/E-FO Model) in grades 3 and 4 and variance explained by grades (R^2) for math ASC, writing ASC, reading ASC, and general ASC

Grade	3		4	
Model	I/E-NMS Model	I/E-FO Model	I/E-NMS Model	I/E-FO Model
Path	β	β	β	β
External comparisons				
Grade Math → gASC	.44***	.44***	.44***	.45***
Grade German →gASC	.28**	.28**	.58***	.59***
Grade Math → mASC	.78***	.84***	.86***	.92***

Grade German → wASC	.37***	.48***	.48***	.72***
Grade German → rASC	.42***	.50***	.49***	.71***
Dimensional comparisons				
Grade Math → wASC	-.40**	-.05	-.56***	-.23**
Grade Math → rASC	-.22**	.09	-.43***	-.16*
Grade German → mASC	-.86***	-.34**	-.92***	-.22**
Interaction grade math and grade German				
Grade Math x German → mASC	.40**	.25**	.13***	.10**
Grade Math x German → wASC	.18*	.14*	.16***	.14***
Grade Math x German → rASC	.20**	.15**	.14***	.14***
R^2				
gASC	.29***	.29***	.46***	.46***
mASC	.41***	.29***	.35***	.36***
wASC	.13*	.15**	.19***	.24***
rASC	.11**	.21***	.14***	.25***

Note. Despite equal nomenclature, mASC = math ASC, wASC = writing ASC, and rASC = reading ASC are first-order correlated factors in the I/E-FO model and specific factors that are residualized by gASC in the I/E-NMS model; thus, factors from the two models are not directly comparable. * $p < .05$ ** $p < .01$ *** $p < .001$.

Dimensional comparisons. In the I/E-NMS model, paths from school grades to the specific ASC in the non-corresponding domain were negative and statistically significant, suggesting contrast effects of dimensional comparisons in both grades (Table 7). In the I/E-FO model, all paths reflecting dimensional comparison processes were negative and significant, except for the paths leading from the math grade to the non-corresponding ASCs in reading and writing in Grade 3. The comparison of both models revealed that paths reflecting external comparisons were slightly larger in the I/E-FO model, whereas paths reflecting dimensional comparisons were larger in the I/E-NMS model. Moreover, in grade 3, the paths reflecting dimensional comparisons from the math grade to the non-corresponding ASCs (reading and writing) did not reach significance in the I/E-FO model.

Achievement differences: Interaction with school grades. In both the I/E-NMS model and the I/E-FO model, the interaction term between school grades in German and math significantly predicted math ASC, reading ASC, and writing ASC (Table 7). The effect of the interaction term was slightly higher in the I/E-NMS model compared to the I/E-FO model. For both models, the size of the interaction effect was comparable in grade 3 and grade 4. This was demonstrated by a non-significant $\chi 2$ difference test between a model with equality constraints and a model without equality constraints of the effects of the interaction between grades 3 and 4 (I/ E-NMS model: $\Delta\chi 2$ (3) = 5.002, p = .172; I/E-FO model: $\Delta\chi 2$ (3) = 4.773, p = .189). Figure 10 provides an example of the interaction effect between school grades in German and math for math ASC within the I/E-NMS model in grade 3.

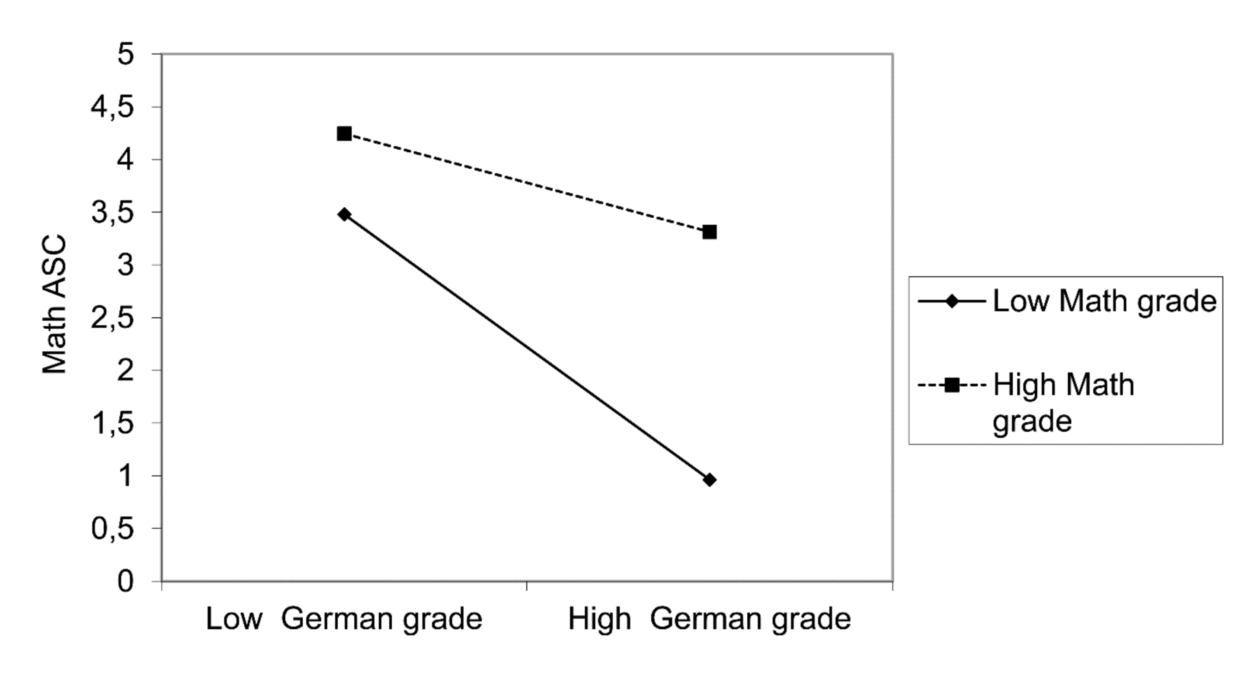


Figure 10. Example: Effects of the interaction term (math × German grade) on academic self-concept in math in grade 3. The intercept is hypothetical because it cannot be estimated using WLSMV

4.6 Discussion

The aim of the present study was to investigate an important aspect of differentiation of ASC: How do profiles of students' strengths and weaknesses form across the elementary school years, and which factors influence profile formation? A key characteristic of our study was that we juxtaposed the NMS model, which provides a very clear picture of students' ASC profiles, with other structural models of ASC, specifically the often-used first-order correlated factor model. To this end, we drew on data from a quasi-representative sample of German elementary school students spanning grades 1 to 4. Our first step was to model different structural assumptions of ASC. Whereas the multidimensional nature of ASC is well documented in the literature, it is still a matter of debate how one should account for its hierarchical nature, specifically for the assumption of a general ASC at the apex. By testing different structural models of ASC, our study added to existing knowledge. First, we are the first to demonstrate that the NMS model accounts for both a general

ASC and for the multidimensionality of ASC already in elementary school students. Second, for this group of students we showed that verbal ASC can be meaningfully differentiated into reading ASC and writing ASC. We found profile formation to take place early during the first three years of elementary school, indicated by initially positive correlations between math and native language (i.e., reading and writing) ASCs decreasing toward zero (in the first-order correlated factor model) or by changing from positive to negative (in the NMS model). Further, results indicated a profile formation within the native language by the decreasing positive correlations between reading ASC and writing ASC during the first three years of school. Within the I/E model, we then tested whether dimensional comparisons influence profile formation in grades 3 and 4. Overall, our findings supported the assumptions of the I/E model. However, effect sizes differed slightly depending on the structural model used. Expanding the I/E model by an interaction term (math grade × German grade), we showed that achievement differences moderated the effect of dimensional comparisons. With increasing differences between students' grades in math and German, contrast effects of dimensional comparisons became stronger.

4.6.1 Limitations

Our study analyzed cross-sectional data only (i.e., samples of students at different grade levels). However, such data can be quite useful for demonstrating age or grade differences and interindividual differences in developmental processes (Robinson, Schmidt, & Teti, 2005). Further, cross-sectional data have proven fruitful for investigating research questions pertaining to development in general (e.g., Gopnik & Astington, 1988; Marsh & Ayotte, 2003). Considering that only longitudinal data can provide information about age- (or grade-) related changes or interindividual differences in developmental trajectories (e.g., Miller, 1998), our findings cannot be interpreted in terms of actual developmental processes per se. Hence, replicating our findings longitudinally is a promising avenue of future research. Our present focus, however, was on profile formation. Therefore, we investigated dimensional comparisons as a source of information for the different subject-specific ASCs as well as external comparisons when testing the I/E model. However, the formation of subject-specific ASC is influenced by further sources and additional frames of references (Skaalvik & Skaalvik, 2002), for example, intraindividual comparisons over time. It is still unclear how these comparisons impact ASC profile formation and how they interact with each other (see Möller & Marsh, 2013). Hence, for understanding ASC profile formation comprehensively, future research should also consider these factors. The fact that only one achievement indicator was available to predict writing ASC and reading ASC in our test of the I/E model is one further limitation which resulted in a mismatch of predictor (grades) and criterion (ASC) levels of generality. In line with the specificity matching principle (Baumeister et al., 2003), specific constructs should be used to predict specific outcomes, whereas general constructs should be used to predict general outcomes (Brunner et al., 2009). Thus, future studies should consider using separate achievement indicators for reading and writing. This would also allow clarifying whether there is (1) an assimilation effect of dimensional comparison processes between reading and writing achievement and their non-corresponding ASC within the verbal domain or (2) a contrast effect between reading/writing achievement and math ASC. Utilizing separate achievement indicators for reading and writing, Ehm et al. (2014) confirmed the I/E model for reading ASC and math ASC in grade 3 but found no contrast effect between either math achievement and writing ASC or writing achievement and math ASC. However, they observed assimilation effects of dimensional comparison processes between writing ASC and reading ASC. It should be noted, however, that Ehm et al.'s (2014) study cannot be compared directly to ours because they used standardized achievement test scores, and these could influence the results of the I/E model (see Möller et al., 2009). Previous research on the I/E model indicates that the operationalization of achievement (grades vs. standardized achievement test scores) moderates the effects of external and dimensional comparison processes.

Correlations between math and verbal achievement were substantially larger for standardized test scores, whereas the different ASCs were more strongly correlated with achievement in the corresponding domain when grades were used (meta-analysis of the I/E model Möller et al., 2009) (see also Marsh et al., 2014).

4.6.2 Structure and profile formation of academic self-concept in elementary school students

At all grade levels, the NMS model with general ASC at the apex and reading ASC, writing ASC, and math ASC as specific factors nested under general ASC provided a very good fit to the data. That is, the theoretical assumptions that (1) ASC is multidimensional (Marsh/Shavelson model) and that (2) general ASC is the central construct in ASC (Shavelson model), which are incorporated in the NMS model, could be supported for elementary school children, too. A special feature of this model is that the correlations among specific factors (i.e., math ASC, writing ASC, and reading ASC) depict students' ASC profiles, that is, how students subjectively assess their strengths and weaknesses in different school subjects independent of their general level of ASC. From a theoretical stance, this model is plausible: When students compare their perceived ability in various subjects, the general level of ASC should be unimportant for evaluating one's profile of strengths and weaknesses. In contrast to the NMS model, subject-specific ASC factors still contain variability attributable to a general ASC in the first-order correlated factor model of ASC, which makes it more difficult to interpret correlations between subject-specific ASCs in terms of students' self-concept profiles. Hence, we think that the NMS model may be better suited for investigating profile formation of ASC than a first-order correlated factor model of ASC, although the latter demonstrated good fit to our data. Using the NMS model, we identified positive correlations between math and reading/writing ASC in grades 1 and 2, but negative correlations in grades 3 and 4, which is in line with our expectations. This finding suggests that (1) the ASC profile developed with increasing grade level and that (2) profile formation started early. Comparing these results with those for the first-order correlated factor model, where correlations between math and writing/reading ASC were positive at every grade level (but lower in the higher grades as well), changes became more apparent in the NMS model. Findings for the NMS model revealed a shift from positive to negative correlations between math and writing/reading ASC and increasingly negative correlations between these different facets of ASC with increasing grade level. In sum, in line with the differential distinctiveness hypothesis, our findings support the assumption that math ASC and native language ASC become more differentiated with age (Marsh & Ayotte, 2003). In Germany, where our study took place, achievement feedback through explicit grading usually starts in grade 3. However, according to our results, elementary school students' profile formation begins in grade 2, suggesting that explicit grading may not be a necessary condition for profile formation of ASC. One possible reason is that grades are only one source of achievement-related feedback. External feedback from significant others (teachers, parents, peers) may also reinforce profile formation of ASC (see, e.g., Marsh et al., 1998). More specifically, several studies showed that children increasingly use available information like ability feedback from significant others to evaluate their own ability during elementary school (cf., Dweck, 2002; Wigfield & Eccles, 2002). Drawing on the differential distinctiveness hypothesis (Marsh & Ayotte, 2003) we further assumed an integration of the different skill-specific ASCs (i.e., writing and reading) indicated by stable or increasing positive correlations with grade level. It should be noted that the differential distinctiveness hypothesis has not yet been explicitly tested with regard to different skill-specific ASCs. In contrast to expectations, the correlation between writing ASC and reading ASC was significantly lower in grade 3 as compared to grade 2, and only slightly higher (and non-significant) again in grade 4 as compared to grade 3. This correlational pattern indicated increasing differentiation within the verbal ASC profile during the first three years of school. How can this be explained? One possibility is that the assumption of the differential distinctiveness hypothesis (i.e., that self-concepts associated with the same higher-order factor are more closely associated) does not apply to the differentiation of skill-specific ASC. The original differential distinctiveness hypothesis rests on the higher-order factor constellation within the structural model of the entire self-concept (Marsh/Shavelson model). However, this model does not include the different skill-specific ASCs of each subject-specific ASC. Further, we hypothesized stronger integration of children's different skill-specific ASCs, drawing on results from previous studies that showed high correlations between skill-specific ASCs of (native) language ASC for different age groups (Arens & Jansen, 2016; Arens, Yeung, & Hasselhorn, 2013; Lau et al., 1999). However, all of these studies (and our study, too) relied on cross-sectional data, which hampers the interpretation of results as a developmental process. Hence, it is also possible that skill-specific ASC becomes more differentiated with age, but not to the same extent as subject-specific ASC, for example, math ASC and verbal ASC. This seems plausible because students might learn that writing and reading are different skills—for instance, a student might be better at writing essays but might not understand texts very well when reading them. However, it is unlikely that a student could write a good essay without being able to understand a text. Therefore, the correlations between writing ASC and reading ASC might be substantial but not perfect, as indicated by previous studies as well as our study. To conclude, future studies using longitudinal data are needed to clarify differentiation within (native) language ASC.

4.6.3 The extended I/E model and factors moderating dimensional comparisons

Although profile formation was observed even without explicit grading (in grades 1 and 2), our findings for the extended I/E model suggest that grading contributes to this process. For both ASC models, most assumptions of the extended I/E model could be replicated in grades 3 and 4 using teacher-assigned grades and the interaction between them as achievement indicators. All paths representing external comparison processes were positive and significant in both models. Findings for internal (dimensional) comparison processes partly depended on the specific ASC model used. In grade 4, all paths representing internal comparison processes were negative and significant regardless of the model, whereas in grade 3 this was only the case for the NMS model. When the I/E model was investigated with the first-order correlated factor model, the paths leading from math grade to reading ASC and writing ASC failed to reach significance in grade 3. One possible explanation for this discrepancy is the specific feature of the NMS model: Subject-specific ASC factors are residualized from general ASC. In the NMS model, correlations between subject-specific ASC in math and reading/writing were moderately negative, reflecting the finding that students see themselves as either “math” or “verbal” persons (Brunner et al., 2009; see also Marsh & Hau, 2004). Statistically, moderately negative correlations between the different subject-specific ASCs should result in larger negative paths between non-corresponding grades and ASC than when correlations are small but positive, which was the case for the I/E-first-order correlated factor model. When comparing the two I/E models, paths indicating external comparison processes revealed slightly higher coefficients in the first-order correlated factor model than in the NMS model, while path coefficients for internal comparison processes were slightly higher (and negative) in the NMS model. An explanation may be that, again, subject-specific ASC in math, reading, and writing are residualized from general ASC in the I/E-NMS model. We found teacher-assigned grades in math and German to be related to both specific ASC and general ASC. Residualization therefore results in smaller positive path coefficients for external comparisons in the I/E-NMS model [10]. In the I/E-first-order correlated factor model, the effects of internal comparisons are confounded with those of external comparisons referring to general ASC (because subject-specific ASC still contains variance attributed to the general ASC). Hence, contrast effects of internal comparisons were smaller whereas effects of external comparison were slightly stronger in the I/E-first-order correlated factor model. Overall, our study demonstrates that the investigated effects are robust across different structural models of ASC and that the I/E model is generalizable to elementary school students—a set of findings which is well in line with previous research (see Möller et al., 2009; Pinxten et al., 2015). Further, differences in detail suggest that when studying the effects of internal and external comparison processes, the moderating role of the structural models of ASC should be kept in mind. One factor that reinforces contrast effects of dimensional comparisons between subjects pertaining to different domains is the achievement difference between subjects. The higher this discrepancy, the stronger the contrast effects of dimensional comparison processes (Rost et al., 2005; Schilling et al., 2004) which implies a more pronounced profile formation of ASC. However, this factor has been much neglected in I/E model research to date. Our findings suggest that the moderating effect of grade differences in cross-domains generalizes to elementary school students in grades 3 and 4. However, objective achievement differences may not be the only moderators of the effect of dimensional comparisons. Studies with secondary school students have shown, for example, that subjective beliefs, specifically, beliefs in a negative interdependence between math and verbal abilities (Möller, Streblow, & Pohlmann, 2006) or in perceived similarity of school subjects (Helm et al., 2016) may also be an important factor. We therefore recommend that future studies of the I/E model consider objective factors such as achievement differences (e.g., as an interaction term), subjective factors such as beliefs, and their interplay as moderators of dimensional comparisons. In general, further research is needed to test whether findings generalize across subgroups such as gender or students from different school systems.

4.6.4 Conclusions and implications

In sum, the NMS model of ASC (Brunner et al., 2008; Brunner et al., 2009; Brunner et al., 2010) may guide future research on the profile formation of ASC. Our findings indicate that this process starts during the first two to three years of elementary school, even without explicit feedback through teacher-assigned grades. However, grade feedback exacerbates differentiation in grades 3 and 4, especially when grades differ between contrasting domains. Our study showed that the effects of dimensional comparison processes depend both on the grade level under examination and on the chosen structural model of ASC. Both factors should, therefore, be taken into account when studying these effects. Our findings have theoretical, methodological, as well as practical implications for the research community. First, different structural models of ASC are suitable for different research questions. When investigating profile formation, a nested ASC model seems to be preferable over a first-order correlated factor model because this model allows controlling for the level of general ASC in students' profiles. Thus, students' ASC profiles of subjective strengths and weaknesses (in terms of subject-specific ASCs) as well as central aspects of the shape of these profiles (in terms of correlations among subject-specific ASCs) can be depicted. The nested ASC model, as applied in the present study, has the advantage that a general ASC factor is psychometrically well defined by the items directly measuring general ASC. Further, the general ASC factor in this model is relatively independent of the subjects investigated. Hence, the present nested ASC model might be especially well suited in cases where not all subjects that are taught in elementary school are studied, as was the case in our study (see Chen, West, & Sousa, 2006; Eid et al., 2003) (see also Brunner et al., 2009). Alternatively, one may also specify a so-called complete bifactor model that includes a general factor and specific factors for all subject-specific academic self-concept s as well as a specific factor that influences the items measuring general ASC. This model has the advantage that relations of covariates (e.g., gender) to the specific factor of items measuring general academic self-concept can be investigated. Such models often approximate the empirical data even better than the nested-factor model applied in the present study (Arens & Morin, 2016). However, the complete bifactor representation has two disadvantages (see Eid, Geiser, Koch, & Heene, 2017). Compared to the nested-factor model applied in the present study, in the complete bifactor model, the general factor is psychometrically less well defined as it depends more strongly on the subjects investigated. Further, complete bifactor models often tend to show irregular loading patterns or vanishing variances of the specific factors for general measures. Such anomalous results are not well aligned with the underlying theories. Irrespective of whether a complete bifactor model or a nested-factor model is chosen to study academic self-concepts, one should take into account that in these models, a significant share of variance is extracted from subject-specific measures, which renders the interpretation of correlations between subject-specific factors and external variables more complex. Moreover, mean levels of subject-specific ASCs are more complex to interpret in the NMS model (or the complete bifactor model) than in the first-order correlated factor model (but see Brunner et al., 2010; Chen et al., 2006) which makes the latter more attractive when mean level changes of subject-specific ASC are examined. Second, from a practical point of view, ASC profiles (i.e., knowing one's own strengths and weaknesses) are useful for self-evaluations and may thus facilitate decision making in educational settings (Möller & Marsh, 2013). However, this may apply only when a student's ASC profile corresponds to actual individual strengths and weaknesses, but not when subjective factors (e.g., beliefs in a negative interdependence of math and verbal abilities; see Möller, Streblow, & Pohlmann, 2006) reinforce dimensional comparisons between cross-domain subjects. Previous research has indicated that self-concept intervention approaches which focus on enhancing comprehensive and appreciative feedback on students' performance in a particular subject were effective in promoting positive ASC (Niepel, Brunner, & Preckel, 2014; O'Mara, Marsh, Craven, & Debus, 2006). Such interventions may also be especially suitable for promoting the development of ASC profiles that correspond to students' actual strengths and weaknesses. This study has demonstrated that student profiles of ASC start to develop very early in elementary school education. Armed with this knowledge, practitioners should consider implementing interventions designed to promote ASC in students right from the start of their school careers.

5 Study 3

Effects of achievement differences for internal/ external frame of reference model investigations: A test of robustness of findings over diverse student samples [18]

[18] Schmidt, I., Brunner, M., & Preckel, F. (2017). Effects of achievement differences for internal/external frame of reference model investigations: A test of robustness of findings over diverse student samples. *British Journal of Educational Psychology.* Advance online publication. https://doi.org/10.1111/bjep.12198

5.1 Abstract

Background. Achievement in math and achievement in verbal school subjects are more strongly correlated than the respective academic self-concepts. The internal/external frame of reference model (I/E model; Marsh, 1986, Am. Educ. Res. J., 23, 129) explains this finding by social and dimensional comparison processes. We investigated a key assumption of the model that dimensional comparisons mainly depend on the difference in achievement between subjects. We compared correlations between subject-specific self-concepts of groups of elementary and secondary school students with or without achievement differences in the respective subjects.

Aims. The main goals were (1) to show that effects of dimensional comparisons depend to a large degree on the existence of achievement differences between subjects, (2) to demonstrate the generalizability of findings over different grade levels and self-concept scales, and (3) to test a rarely used correlation comparison approach (CCA) for the investigation of I/E model assumptions.

Samples. We analysed eight German elementary and secondary school student samples (grades 3–8) from three independent studies (*N*s 326–878).

Method. Correlations between math and German self-concepts of students with identical grades in the respective subjects were compared with the correlation of self-concepts of students having different grades using Fisher's Z test for independent samples.

Results. In all samples, correlations between math self-concept and German self-concept were higher for students having identical grades than for students having different grades. Differences in median correlations had small effect sizes for elementary school students and moderate effect sizes for secondary school students.

Conclusions. Findings generalized over grades and indicated a developmental aspect in self-concept formation. The CCA complements investigations within I/E-research.

5.2 Theoretical background

Self-concepts comprise subjective beliefs about the qualities characterizing a person, including their attributes and beliefs about who and what the self is (e.g., Baumeister, 1999). Self-concepts can be conceptualized as multidimensional and hierarchically structured. The prominent Shavelson model (Shavelson et al., 1976) puts a general self at the most general level of self-concept and four broad domains of self- concept (academic, social, physical, and emotional) at a subordinate level. In this study, we focused on academic self-concepts, which refer to students' personal beliefs about their academic abilities (e.g., Trautwein, Lüdtke, Marsh, & Nagy, 2009). These self-concepts are important in educational psychology because they influence, for example, decision-making (e.g., course choices; Nagengast & Marsh, 2012) and interest (e.g., Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005), and are reciprocally related to achievement (reciprocal effects model, REM; e.g., Marsh & Martin, 2011).

In different theories and models of self-concept, comparison processes within different frames of references are described as important sources for self-concept development (temporal comparison theory:Albert, 1977; expectancy-value theory: Eccles, 2009; e.g., social comparison theory: Festinger, 1954). For example, students compare their ability in a subject across time (temporal comparisons) and with specific students or groups of students like their class (social comparisons). A well-replicated effect of social comparisons with a group of students (e.g., class) is the big-fish–little-pond effect (BFLPE; e.g., Marsh, 2005). The BFLPE posits that equally achieving students have higher self-concepts when they are in a lower achieving group than when they are in a higher achieving group. In general, comparison processes support self-evaluations but also self-enhancement, self-maintenance, and self-improvement (e.g., Möller & Marsh, 2013).

Recent research on self-concepts focuses on dimensional comparisons. In dimensional comparisons, students contrast their own achievement in one subject/domain with their own achievement in another subject/domain. These processes serve as one explanation for the finding that academic self-concepts are multidimensional with separate, domain-specific self-concepts for various achievement domains such as specific school subjects (e.g., Brunner et al., 2010). Especially, math self-concepts and verbal (native language) self-concepts were found to be nearly uncorrelated or only weakly correlated (e.g., Marsh, 1990b, 1990c). This finding strongly contrasts the ubiquitous finding that those students with strong (weak) achievement in math typically also demonstrate strong (weak) verbal achievement – a correlational pattern among achievement measures that suggests substantial correlations between corresponding self-concepts (Marsh, 1986, 1990b, 1990c). To explain these seemingly paradoxical correlational patterns, the I/E model (Marsh, 1986) was developed. According to this model, the self-concept in a particular school subject is mainly formed in relation to two frames of reference. Within an external frame of reference, students engage in social comparisons; that is, they compare their own achievement in a particular school subject with the achievement level of the peer group (e.g., class-average achievement) and related information such as grade distribution in that subject (Möller et al., 2009). If students' subject-specific achievements are higher than the class-average achievement, the subject-specific self-concept will also be higher (Möller et al., 2009, p.1131). As achievements between subjects are usually substantially positively correlated, social comparisons would predict that the corresponding subject-specific self-concepts are substantially correlated, too.

Within an internal frame of reference, students engage in dimensional comparisons (Marsh et al., 2014; Möller & Marsh, 2013); that is, they compare their own achievement in one particular school subject with their own achievement in other school subjects (Marsh, 1986). Thus, the *difference* between an individual student's achievement in one subject and the achievement of the very same student in another subject drives the development of self-concepts (Marsh, 1986). This means that students may have a high math self-concept even if they are not achieving well in this subject relative to other students (social comparisons) because they are better in this subject than in the native language (dimensional comparisons) (e.g., Marsh, Xu, & Martin, 2012). The combined effect of dimensional and social comparisons leads to a lower correlation between math and verbal self-concepts than between the respective subject-specific achievements (Marsh, 1986).

In general, dimensional comparisons can either reduce or increase correlations between subject-specific self-concepts. According to the dimensional comparison theory (Helm et al., 2016; Marsh et al., 2014; Möller & Marsh, 2013), lower correlations between two subject-specific self-concepts are more likely when subjects are perceived as dissimilar (contrast effect), for example, math and native language. Higher correlations between subject-specific self-concepts are expected if the subjects are perceived as similar (assimilation effect), for example, math and physics (Jansen et al., 2015).

There are different approaches to test the effects of social and dimensional comparisons on subject-specific self-concepts as stated by the I/E model. The majority of studies applied a path-analytic approach, in which math self-concept is regressed on math and verbal achievement and verbal self-concept is regressed on math and verbal achievement (e.g., Marsh, 1986). Social comparison processes imply positive paths from math achievement to math self-concept and from verbal achievement to verbal self-concept reflecting that a student's higher achievement in one subject relative to his or her peers is related to a higher level of self-concept in the same subject. Dimensional comparison processes, on the other hand, imply negative paths from math achievement to verbal self-concept and from verbal achievement to math self-concept reflecting contrast effects of an improvement in achievement in one subject on the self-concept in the other subject. Importantly, in the path-analytic approach, dimensional comparison effects are specified such that achievement in one subject is controlled for (i.e., held constant) while achievement in the other subject may vary. In doing so, it is possible to relate achievement differences between math and verbal subjects (for the very same student) to the subject-specific self-concepts. Within this approach, there is broad empirical support for the predictions of the I/E model; findings generalize over different measures of achievement or self-concepts as well as over gender and country (e.g., Möller et al., 2009).

Concerning the generalization of the findings to different age groups, a few studies indicated effects of dimensional comparisons from grade 3 onwards (Möller et al., 2009). However, previous research on the I/E model (irrespective of the applied statistical approach) focused primarily on secondary school students, whereas studies with elementary school students are relatively scarce (e.g., Pinxten et al., 2015). For this age group, there is some evidence that the effects of dimensional comparisons are somewhat lower; that is, the correlations between math and verbal academic self-concept are higher in elementary school students than in secondary school students. In a meta-analysis of studies investigating the I/E model (Möller et al., 2009), correlations between math and native language self-concept were slightly lower in students who spend more years in school. However, only three of the 69 samples consisted of elementary school students. Thus, it is difficult to generalize these findings. Developmental and environmental aspects can explain the lower differentiation of self-concept in younger students. In particular, the cognitive development of children relates to their ability to understand, interpret, and integrate ability-related experiences such as feedback from significant others (e.g., teachers, parents, peers) and information from social comparisons into one's self-concept (e.g., Wigfield & Eccles, 2002). Further, dimensional comparisons become more differentiated over time (Harter, 1999). Hence, self-concept becomes more differentiated with age and feedback is more integrated into students' self during elementary school (e.g., Harter, 1999). Investigating factors influencing self-concept formation of elementary school students is important because already in this age group self-concepts influence achievement and vice versa (e.g., Guay et al., 2003). Therefore, research on the I/E model in elementary school students may be practically relevant for teachers' feedback practices and for interventions aiming to promote self-concepts.

A further approach to test the I/E model[19] applied the logic of experimental control techniques to separate the effects of dimensional and social comparisons (Rost et al., 2005; Schilling et al., 2004) We term this method correlation comparison approach (CCA).[20] This approach tests the key assumption of the I/E model that dimensional comparisons depend on the existence of achievement differences between subjects. To do so, a student sample is split into groups. One group comprises students who show the same level of achievement (e.g., same grades) in, for example, math and verbal subjects. For this group, the influence of dimensional comparisons is controlled for because comparisons between subjects‘ provides no information, or information that can be ignored' (Rost et al., 2005, p. 558) with regard to students' strengths and weaknesses. Thus, only social comparisons should affect self-concepts in this group of students. The other group comprises students who show different achievements in math and verbal subjects. For this group, dimensional as well as social comparison processes should affect math and verbal self-concepts. Given that dimensional and social comparison processes work in opposite directions, a higher correlation between math and verbal self-concepts is expected in the group of students with identical achievement in both subjects than in the group of students with different achievements (Rost et al., 2005). The difference in the correlation of self-concepts between both groups depicts the effects of dimensional comparisons. Using the CCA, Rost et al. (2005) could support the predictions of the I/E model in a sample of seventh and eighth graders from the middle and the top track of the German

[19] Skaalvik and Rankin (1992) had proposed the logic of the approach. However, their empirical procedure was somewhat problematic (for a discussion, see Möller et al., 2002).

[20] Rost et al. (2005) used the term quasi-experimental. However, the use of this term is somewhat unusual because they did not actively manipulate the independent variable, which is a necessary precondition to denote a design as experimental (e.g., Shadish, Cook, & Campbell, 2002).

school system. The correlation between math self-concept and German self-concept was substantial and higher for students having equal grades in these subjects (r = .42) than for those having unequal grades (r = .02). However, to our knowledge this is the only study that used this approach for I/E model testing. What are advantages of the CCA relative to the path-analytic approach? First, and perhaps most importantly, the CCA directly tests the key assumption of the I/E model that dimensional comparisons depend on the existence of achievement differences. In a path-analytic approach, this assumption is indirectly incorporated in a model-based manner. Second, the CCA responds to the principle of a minimally sufficient analysis as proposed by the Task Force on Statistical Inference of the American Psychological Association (Wilkinson & Task Force on Statistical Inference, American Psychological Association, Science Directorate, 1999) which states that ‘if the assumptions and strength of a simpler method are reasonable for your data and research problem, use it’ (Wilkinson & Task Force on Statistical Inference, American Psychological Association, Science Directorate, 1999, p. 598). Specifically, the computation of (manifest) correlation coefficients as well as their comparison across groups is easy to implement and operates on available data (see Rost et al., 2005). Path analysis, on the other hand, uses quite complex statistical methods that require estimating unknown parameters. Third, in contrast to the CCA, path analysis may suffer from (multi)collinearity among the achievement measures because these measures are often highly correlated resulting in instable parameter estimates, increasing standard errors, and lower statistical power to assess effects (Cohen, Cohen, West, & Aiken, 2003). To sum up, main evidence for the I/E model comes from studies using path analysis, whereas the CCA is rarely used.3[21] In addition, research concerning elementary school students is still scarce compared to findings for secondary school students.

[21] Several studies investigated the I/E model with other methods such as experimental studies (e.g., Strickhouser and Zell, 2015) or introspective studies (e.g., Möller and Husemann, 2006)

5.3 The present study

This study had four research aims: (1) We tested the key assumption of the I/E model that dimensional comparison processes depend on achievement differences; this assumption is not directly addressed in the path-analytic approach which dominated previous research. Thus, this study contributes to existing knowledge. (2) Most previous research on the I/E model relied on data from students in secondary schools. We examined the generalizability of the I/E model over diverse age groups and included elementary school students. (3) Further, we tested the robustness of findings for different self-concept scales. (4) We analysed data from student samples (grades 3–8) by means of a simple method, the CCA, that was rarely used in previous research. We examined student samples from three independent studies including elementary and secondary schools. We divided each student sample into two groups: one group of students with identical achievement (i.e., grades) in math and German and one group of students with different achievements in these subjects. For students having identical grades, dimensional comparisons are not or less informative for their profile of strengths and weaknesses and social comparisons should mainly affect their self-concepts (see Rost et al., 2005). Thus, the correlation between math and German self-concepts in this group of students mainly reflects social comparison processes: Stronger (weaker) achievement in math/German is expected to coincide with a higher (lower) level of self-concept in math/German, respectively. Given that math and verbal achievement are positively correlated, we expect a positive correlation between math and German self-concepts in this student group. For students having different grades, both dimensional and social comparisons affect their self-concepts. The correlation between math and German self-concepts, thus, reflects the combined effect of both comparison processes. Because dimensional comparisons lower the correlation between math and German self-concepts, the correlation in this student group is expected to be lower than in the other student group. Consequently, the difference between the correlations of math and German self-concepts as obtained for the two groups of students reflects the effect of dimensional comparison processes. Self-concept formation depends on developmental and environmental aspects (e.g., Harter, 1999). Therefore, the effects of dimensional comparisons may be somewhat lower in elementary school students compared with secondary school students. To tackle our research objectives, we implemented the CCA in a manifest variable framework and a latent variable framework. In this study, we focus on the results obtained in the manifest variable framework because it meets the principle of a minimally sufficient analysis (Wilkinson, L. & Task Force on Statistical Inference, American Psychological Association, Science Directorate, 1999). Nonetheless, we replicated our results within a latent variable framework to control for the proportion of variance in subject-specific self-concepts attributable to measurement error and method effects. Thus, relative to the manifest variable framework, the latent variable approach enabled us to compare findings over studies with higher methodological rigour.

5.4 Method

5.4.1 Studies and measures

Study 1: Elementary school (grades 3 and 4) Study 1 (S1) comprised two quasi-representative cross-sectional samples of German elementary school students. This investigation included students in grade 3 (S1/3; N = 326; 52.5% female; age: M = 8.45 years, SD = 0.58) and in grade 4 (S1/4; N = 808; 51.9% female; age: M = 9.56 years, SD = 0.58). Writing self-concept, reading self-concept, and math self-concept were assessed with 3-item scales each from the FEESS-K (Baudson & Preckel, 2015). The items included ‘I am good at reading/writing/math’, ‘I do well in reading/writing/math’, and ‘In reading/writing/math I do most things well’. Students answered on a 3-point rating scale with frowning, neutral, and smiling ‘faces’ to indicate disagreement, indifference, and agreement. Cronbach’s alpha values were good (grade 3: reading: α = .81, writing: a = .82, math α = .86; grade 4: reading: α = .86, writing: α = .87, math: α = .86). Parent reported their children’s grades at the time of data collection (i.e., from their most recent report card).

Study 2: German secondary school (grades 5, 6, and 8) Data of Study 2 (S2) stem from an ongoing longitudinal study in which students in two federal states of Germany (Rhineland-Palatinate and Bavaria) of the top track of the German secondary school system (Gymnasium) participated. The study started in 2005 and assessed students three times in grade 5 and once in grades 6, 8, 11, and 13, respectively. In this investigation, we used data from students in Rhineland-Palatinate who were in the middle of grade 5 (S 2/5; N = 878; 50.3% female; age: M = 10.51 years, SD = 0.56), in the middle of grade 6 (S 2/6; N = 838; 50.8% female; age: M = 11.54 years, SD = 0.55), and in the middle of grade 8 (S 2/8; N = 841; 51.8% female; age: M = 13.5 years, SD = 0.57). Math self-concept and German self-concept were captured with five items of the German translation of the Self-Description Questionnaire by Marsh (1990a; e.g., ‘Math/German is one of my best subjects’). Students responded to these items on a 5-point rating scale ranging from 1 (strongly disagree) to 5 (strongly agree), with higher scores indicating higher self-concepts. Cronbach’s alpha values were good (grade 5: math: α = .88, German: α = .86; grade 6: math: α = .90, German: α = .88; grade 8: math: α = .91, German: α = .88). Students self-reported their school grades from midterm report cards in grades 5, 6, and 8.

Study 3: German secondary school (grades 5, 6, and 7) Data of Study 3 (S3) stem from another longitudinal project. This project studied German secondary school students from two federal states of Germany (Bavaria and Baden-Wuerttemberg) in the top track of the German secondary school system (Gymnasium). Students participated at the beginning of grade 5, at the end of grade 5, at the end of grade 6, and in the middle of grade 7. For this investigation, we used data from the end of grade 5 (S3/5; $N = 499$; 44.9% female; age: $M = 10.9$ years, $SD = 0.47$), the end of grade 6 (S3/6; $N = 461$; 46% female; age: $M = 11.86$ years, $SD = 0.47$), and the middle of grade 7 (S 3/7; $N = 409$; 44.3% female; age: $M = 12.62$ years, $SD = 0.55$). German self-concept was assessed with three items and math self-concept was assessed with four items of the German translation of the Self-Description Questionnaire by Marsh (1990a; e.g., "Math/ German is one of my best subjects"). Students responded to these items on a 5-point rating scale ranging from 1 (strongly disagree) to 5 (strongly agree), with higher scores indicating higher self-concepts. Cronbach's alpha values were acceptable to good (grade 5: math: $\alpha = .88$, German: $\alpha = .77$, grade 6: math: $\alpha = .89$, German: $\alpha = .71$, grade 7: math: $\alpha = .89$, German: $\alpha = .77$).

School grades were obtained from report cards in the end of grades 5 and 6 and in the middle of grade 7. School grades ranged from 1 (very good) to 6 (insufficient). In all studies, student participation was anonymous and voluntary. Written parental consent was obtained for all participating students. Data collection of self-concepts took place in group sessions in classes.

5.4.2 Data analysis

At first, we divided each of the eight samples (i.e., S1/3, S1/4, S2/5, S2/6, S2/8, S3/5, S3/6, S3/7) into two groups: one group of students with identical grades in math and German and the other group of students with different grades in the corresponding subjects. Prior to the main analyses, we wanted to make sure that it is insignificant which grades were equal or unequal (e.g., for our analyses, it did not matter whether students had a "2" [good] in math and German or a "4" [passed] in math and German). We examined whether group membership was associated with achieved grade level or self-concepts by calculating squared correlations between group membership, school grades, and self-concepts for each study and grade level. The degree of association was assumed negligible if the squared correlations (R^2) were smaller than .09, indicating that <9% (representing a small effect) of the variance was explained by group membership (Ellis, 2010). For the hypothesis tests within the manifest variable framework, we used the statistical software SPSS 23 (IBM Corp., 2014). We calculated correlations between German self-concept and math self-concept in the two student groups. Within the latent variable framework, correlation coefficients were estimated by means of confirmatory factor analysis with the software Mplus 8 (Muthén & Muthén, 1998-2017). Math self-concept and German self-concept were specified as correlated latent variables separately for each of the two groups in a multi-group comparison. Scales showed metric measurement invariance over groups as a precondition to compare correlations over groups (see Appendix C for the procedure/results). All items measuring a subject-specific self-concept loaded only on the corresponding latent factor. To account for method effects due to the same wording of items in Study 1, the residual terms of the items ("In reading/math I do most things well", "I am good in reading/writing") were allowed to correlate. The residual terms of the items ("In math/German I learn fast") in Study 2 and the residual terms of the items ('I learn fast in math/German', 'I am hopeless in math/German') in Study 3 were allowed to correlate. To obtain unbiased test statistics, we accounted for the fact that students' were nested within classes using the "type = complex" option in Mplus in combination with a maximum-likelihood estimator which is robust against mild violations of normality. To deal with missing data,[22] we used full-information maximum-likelihood estimation (Enders, 2010).

To evaluate the model fit of the confirmatory self-concept models, we used the chi square goodness-of-fit statistic and the following descriptive fit indices (Hu & Bentler, 1999): Root mean square error of approximation should be below .08, the comparative fit index should be above .95, and the standardized root mean square residual should be below .08.

Within the manifest as well as the latent variable framework, Fisher's Z test for independent samples with one-tailed testing was used to compare correlations between groups by inferential

[22] Covariance coverage across all samples (range: 95–100%).

statistics (for the formula, see Cohen & Cohen, 1983, p.54). Cohen's q (1988, p.109) was used to quantify the size of the effect for the correlation differences between groups with $q \geq .1$ indicating a small effect, $q \geq .3$ indicating a moderate effect, and $q \geq .5$ indicating a large effect.

5.5 Results

5.5.1 Preliminary analyses

Group membership (i.e., students having unequal vs. equal grades) was to a negligible extent related to self-concepts or grades (see Table 8) indicated by the shared variances (R^2s) which were considerably lower than the cut-off value of .09 (R^2 range: 0.00 to 0.03). Thus, group membership did not depend on which grades were equal or unequal. Further, for the confirmatory factor analyses, in all samples the model fit of the self-concept models in the two groups was acceptable to good (see Tables C1 and C2).

Table 8. Rank-biserial correlations between group membership (equal vs. unequal grades) with math academic self-concept (MASC), German academic self-concept (GASC) and the grades in math (M) and German (G) respectively

Study/ grade level	N	r MASC	R^2 MASC	r GASC	R^2 GASC	r M	R^2 M	r G	R^2 G
S 1/3	321	-.020	.00	R .052	.00	-.023	.00	.075**	.01
				W .071	.01				
S 1/4	791	.039	.00	R -.037	.00	.038	.00	-.111**	.01
				W -.106**	.01				
S 2/5	856	.031	.00	.102**	.01	-.022	.00	-.127**	.02
S 2/6	810	.100**	.01	-.003	.00	.088*	.01	-.029	.01
S 2/8	813	.135**	.02	-.105**	.01	-.145**	.02	.057	.01
S 3/5	499	-.026	.00	.057	.00	-.048	.00	.048	.00

S 3/6	461	.079	.01	-.035	.00	.138**	.02	-.018	.00
S 3/7	408	.081	.01	-.107*	.01	.161**	.03	-.086	.01

Note. Correlations were calculated using SPSS with listwise deletion. R=Reading self-concept; W=Writing self-concept. Group membership was dummy-coded (0=unequal grades; 1=equal grades). ****p*<.001** *p*<.01 * *p*<.05.

5.5.2 Hypothesis testing

Table 9 depicts the manifest as well as the latent correlations between math self-concept and German self-concept in the two groups of students with equal or unequal grades and the total samples. In all total samples, the correlations between the grades in math and German were consistently higher than the correlations between the respective self-concepts. This correlational pattern was found for elementary as well as secondary school students. Further, both analytic approaches (i.e., manifest and latent) revealed significantly higher correlations of math and German self-concepts for students having equal grades than for students having unequal grades. This finding applied to all samples apart from S1/3. Median effect sizes for the differences of the correlations (see Table 9) were low for elementary school students (manifest approach: q = .24; latent approach: q = .27), moderate for secondary school students within the manifest approach (q = .43), and high within the latent approach (q = .53). Effect sizes ranged from no effect in S1/3 (manifest: q = .05; latent: q = .13) to a high effect in S2/5 (manifest: q = .74; latent: q = .84).

Table 9. Manifest and latent correlations of math academic self-concept (MASC) and German academic self-concept (GASC) for total student samples, for students with equal grades in math (M) and German (G), and for students with unequal grades in math and German

	Total sample		Equal grades		Unequal grades				
	M with G	GASC with MASC	GASC with MASC		GASC with MASC		Fisher's Z test		Cohen's *q*
Study/grade	*r*	*r*	*N*	*r*	*N*	*r*	*z*	*p*	*q*
Manifest variable approach									
1/3	.68	R .16	143	R .19	178	R .14	.453	.325	.05
		W .18		W .31		W .09	1.945	.026	.23
1/4	.64	R .13	293	R .30	498	R .05	3.509	<.001	.26
		W.15		W.31		W.07	3.387	<.001	.24
2/5	.37	.15	402	.55	466	-.10	10.518	<.001	.74
2/6	.48	.08	372	.36	453	-.09	6.545	<.001	.47
2/8	.52	-.01	339	.35	484	-.18	7.689	<.001	.55
3/5	.54	.06	249	.28	250	-.10	4.308	<.001	.39
3/6	.51	.07	234	.29	227	-.07	3.932	<.001	.37
3/7	.37	.10	140	.25	268	.06	1.856	.032	.20
Latent variable approach									
1/3	.64	R.13	146	R.21	180	R .08	1.176	.120	.13
		W.14		W.30		W.04	2.383	.009	.27
1/4	.63	R .13	302	R .33	506	R .04	4.193	<.001	.31
		W.14		W.31		W.05	3.729	<.001	.27
2/5	.38	.15	406	.60	472	-.15	12.492	<.001	.84
2/6	.49	.10	378	.43	460	-.12	8.281	<.001	.58
2/8	.52	.00	349	.43	492	-.19	9.276	<.001	.65
3/5	.54	.05	249	.33	250	-.15	5.484	<.001	.49
3/6	.51	.06	234	.34	227	-.13	5.170	<.001	.49
3/7	.37	.09	140	.23	269	.05	1.751	.04	.18

Note. R = Reading self-concept; W = Writing self-concept. Inferential statistical test is based on Fisher's *Z differences test* between the two groups. Effect size measure for the difference between the correlations in the two groups was calculated with Cohen's *q*.

5.6 Discussion

In the present study, we tested the key assumption of the I/E model that effects of dimensional comparisons depend to a large degree on the existence of achievement differences between subjects. Further, we tested the generalizability of findings over different grade levels and self-concept scales using a CCA. We found evidence for the assumption that dimensional comparison processes depend on achievement differences and for the generalizability of findings over elementary and secondary school students as well as over different self-concept scales. Overall, our findings supported the usefulness of the CCA, which has rarely been applied in previous I/E model research.

5.6.1 Limitations

Despite the strengths of our study, there are several limitations. First, we only used student samples from Germany. Further, the samples of secondary school students stem solely from the top track of the German multipartite school system. However, on the basis of previous I/E model research (e.g., Möller et al., 2009), we would expect that our findings generalize across the different tracks of the German school system (e.g., Möller et al., 2009). In addition, we focused only two subjects (math and German), for which a contrast effect of dimensional comparisons is assumed (e.g., Marsh et al., 2014). Nevertheless, as stated by dimensional comparison theory (e.g., Helm et al., 2016), dimensional comparisons could result in either contrast or assimilation effects (e.g., assimilation for math and physics; Jansen et al., 2015) according to the (perceived) similarity between subjects.

5.6.2 General discussion

Using the CCA, our study showed that the effects of dimensional comparisons are stronger when achievement differs between subjects. Therefore, our findings are in support of a key assumption of the I/E model. In contrast to popular path-analytic approach, this assumption is directly tested within the CCA. Specifically, we found higher correlations between self-concepts in the group of students with equal grades than in the group of students with different grades. This correlational pattern depicts the combined effects of dimensional and social comparison processes on self-concepts as stated by the I/E model (see Marsh, 1986).

We found a comparable correlational pattern for elementary school students and for secondary school students. Thus, the effects of dimensional comparisons generalized across ages. However, the median effect size for the correlational difference between the group of students having equal grades and those students with different grades was lower in elementary school than in secondary school samples. This indicates that with increasing age, students more strongly perceive a distinct profile of strengths and weaknesses. This interpretation aligns well with previous findings (e.g., Möller et al., 2009) and with developmental theories, which state that self-concept becomes more differentiated with age (e.g., Harter, 1999). However, it remains an open question whether only dimensional comparisons cause this effect. Other factors such as students' beliefs about the (in) dependence of math and verbal abilities (see Möller, Streblow, Pohlmann, & Köller, 2006) or related beliefs of significant others (e.g., teachers) might be important, too. Furthermore, the formation of a profile of strengths and weaknesses over time is likely caused by reciprocally negative effects of dimensional comparisons across subjects (see evidence for the reciprocal I/E model; e.g., Möller, Retelsdorf, Köller, & Marsh, 2011).

5.6.3 Implications

Teacher-assigned achievement-related feedback is important for students' self-concept development (e.g., Burnett, 2003). Our finding that students' subject-specific self-concept is affected by dimensional comparisons, especially when grades between subjects differ, is practically relevant for teachers' feedback practices. It may give teachers an explanation as to why students choose a particular subject, for example, why even a high-achieving student in math might choose a language major. Relative to her math achievement, her language achievement might be even better; accordingly, her subjective strength lies in the verbal field and not in math. The effects of dimensional comparisons further explain why students have a relatively high self-concept in a subject even if they are not achieving especially well in this subject compared to other students. These students are better in this subject than in another (contrasted) subject (e.g., Marsh et al., 2012). In such cases, it is even more important that the teacher is aware of the effect of dimensional comparisons and still relies on the objective performance in a specific subject when giving feedback, because accurate and comprehensive achievement-related feedback has been found to be most effective in a meta-analysis on interventions to enhance self-concept (O'Mara et al., 2006). As a methodological implication, we found that these complements methods for I/E studies. Depending on the respective aims and data, researchers can exploit the advantages of either the path-analytic approach or the CCA. The CCA is advantageous if researchers aim to directly test the key assumption of the I/E model that dimensional comparisons hinge on achievement differences between subjects; in path-analytic approaches, this assumption is indirectly incorporated in a model-based manner. Further, the CCA is easy to implement and operates on available data; therefore, it responds to the principle of a minimally sufficient analysis (Wilkinson, L. & Task Force on Statistical Inference, American Psychological Association, Science Directorate, 1999). Moreover, the approach is not prone to possible statistical problems due to (multi)collinearity regarding the correlation between subject-specific achievement indicators. In contrast, the path-analytic approach is advantageous if researchers want to quantify also the effect of social comparisons. Further, it offers a more flexible tool to investigate moderators and/or mediators within the I/E model.

Aside from practical or methodical implication regarding the I/E model and the dimensional comparison theory, there are some theoretical implications of our findings (see Möller & Marsh, 2013). These concern extensions of the I/E model to further school subjects to obtain a broader picture of students' profile of strengths and weaknesses but also extensions to other constructs (e.g., interests) in the sense of a generalized I/E model (Möller, 2016) Further, to subtilize dimensional comparison theory, the integration of other models such as the inclusion/exclusion model (Schwarz & Bless, 1992) in the dimensional comparison theory may be promising. This model offers another perspective with regard to the circumstances/contextual factors under which spontaneous assimilation or contrast effects of comparisons may occur. In summary, by showing that our findings could be replicated across different age groups within the CCA, we endorse this method for lacked extensions and theoretical refinements regarding dimensional comparisons theory based on the I/E model.

5.7 Acknowledgements

Data of Study 1 stem from a project by Franzis Preckel and Tanja Gabriele Baudson on intelligence and cognitive motivation in German elementary schools. The project was funded by the Hogrefe Publishing Company, Göttingen. Data of Study 2 stem from a longitudinal study conducted by Franzis Preckel which focuses on motivational and self-concept development in secondary school students. This longitudinal study is supported by the Ministry of Education, Science, Adolescence, and Culture of Rhineland-Palatinate. Data of Study 3 stem from the PULSS Project conducted by Wolfgang Schneider, Eva Stumpf, Franzis Preckel, and Albert Ziegler, a longitudinal study on school Achievement and academic development at the beginning of secondary education. The project was funded by the Bavarian Ministry for Education and Culture, by the Ministry for Culture, Youth, and Sports of the Federal State of Baden-Wuerttemberg, and by the Karg-Foundation, Frankfurt/Main, Germany. Funding sources of all studies were not involved in decisions referring to study design or analysis or interpretation of the data.

6 Discussion

This discussion aims to summarize the major findings on the three areas of self-concept (self-esteem, academic self-concepts, and social self-concepts) investigated in samples of students in childhood and adolescence and to reflect on the central findings with respect to the research aims. In addition, theoretical and practical implications derived from the present findings will be outlined, followed by a consideration of their strengths and limitations as well as an outlook for further research. Finally, overarching conclusions from the three studies included in this dissertation will be drawn.

6.1 Summary of results and reflection of central findings

In the following, the main findings from the three studies are summarized, followed by a reflection of the central findings with respect to the structure and the development of self-concepts and their determinants.

6.1.1 Summary

Results from Study 1 indicated that self-esteem, general academic self-concept, social self-concept of acceptance, and social self-concept of assertion became more stable in rank order over time but differed in mean-level change and the shape of mean-level changes from late childhood to late adolescence (grades 5 to 11; German students from the top track). In particular, self-esteem stayed constant, general academic self-concept showed a mean-level decline, and the social self-concept of acceptance and assertion increased. Shapes of mean-level change were nonlinear for social self-concept of assertion and academic self-concept and linear for social self-concept of acceptance. There were remarkable interindividual differences in the change in all self-concept facets. Further, some longitudinal causal relations between self-concept facets mainly occurred between grades 6 and 8 (t4-t5). Bottom-up effects, top-down effects, and reciprocal effects were observed. Pattern of bottom-up effects and top-down effects were relatively balanced meaning that a comparable amount of top-down and bottom-up effects were observed across the time span; the evidence for reciprocal effects (i.e., bottom-up and top-down at the same time) was weak. There was also evidence for transdimensional effects between academic self-concept and social self-concepts but no effects were found between both types of social self-concepts.

Results from Study 2 demonstrated that the nested Marsh/Shavelson model of academic self-concept accounts for both a general academic self-concept at the apex of the hierarchy and for the multidimensionality of academic self-concept (math, reading, and writing) in four quasi-representative samples of German elementary school students in grades 1 to 4. Further, students' verbal academic self-concept can be differentiated into reading academic self-concept and writing academic self-concept. Structural change (i.e., differentiation and integration) took place early during the first three years of elementary school, indicated by initially positive correlations between math and native language (i.e., reading and writing) academic self-concepts decreasing toward zero (in the first-order correlated factor model) or by changing from positive to negative (in the nested Marsh/Shavelson model). In addition, results indicated a differentiation within the native language by the decreasing positive correlations between reading academic self-concept and writing academic self-concept during the first three years of elementary school. The findings supported the assumptions that dimensional comparisons influence structural changes (i.e., profile formation) in grades 3 and 4 within the I/E model. However, effect sizes differed slightly depending on the structural model used. Moreover, the grade difference between math and native language moderated the effect of dimensional comparisons. In particular, with increasing differences between students' grades in math and native language, contrast effects of dimensional comparisons became stronger.

Results from Study 3 support evidence for the key assumption of the I/E model that dimensional comparison processes depend on achievement differences in samples of German elementary school students and secondary school students (grades 3 to 8) investigated with a seldom-used statistical approach (i.e., the correlational comparison approach[23]). Findings of the I/E model generalize across findings from elementary and secondary school students as well as across different self-concept scales. However, the effects of dimensional comparisons were stronger in elementary school students than in secondary school students.

6.1.2 Reflection of central research findings regarding aims

6.1.2.1 Structure of academic self-concept and its structural change and determinants

Since the seminal work of Shavelson et al. (1976) in which they propose that the self-concept is multidimensional and hierarchically structured construct, a large number of subsequent studies investigated the proposed structural assumptions. Measurement instruments were generated to investigate multidimensional (e.g., SDQs, ASDQs; see Section 2.1 for a description) and a vast amount of studies found evidence for the multidimensionality of self-concept (cf., Marsh et al., 2012). That is self-concept facets are separate but correlated.

However, the hierarchical aspect of self-concept is a matter of debate. Particularly a) how to model the hierarchical aspect and b) how to place general self-concept assessed with a unidimensional scale in the structural model of self-concept in a factor model (see Arens & Morin, 2016). This topic concerns the entire self-concept but as well as self-concept areas (academic and non-academic self-concept (i.e., emotional, social and physical) (see the Shavelson et al. model of self-concept in Figure 1).

[23] Note, we renamed the approach (original: quasi-experimental approach; Rost et al., 2005).

Note that the original Shavelson et al. model is depicted as a higher-order factor model. General self-concept, general academic self-concept, or general nonacademic self-concepts assessed with a unidimensional scale are not included. Stated differently, general self-concept (i.e., self-esteem) and general academic self-concept is depicted as a higher-order factor in the Shavelson et al. model (see Figure 1). Because several studies found evidence that math and reading self-concepts are nearly unrelated, challenging the assumption of a higher-order factor in the academic section of the self-concept, the Shavelson et al. model was revised by Marsh and Shavelson (1985) to show the strong domain specificity of the academic self-concept. In the Marsh/Shavelson model, academic self-concept is depicted as a higher-order factor model with two independent higher-order factors at the apex—a verbal and math domain general academic self-concept. In this model general academic self-concept assessed with a unidimensional scale is a subordinate as well as subject-specific self-concept. However, several studies identified positive correlations of general academic self-concept (assessed with a unidimensional scale) with both native language and math academic self-concept (e.g., Marsh, 1986; Marsh et al., 1988). This and other critical findings (see Section 2.1) indicate that native language and math academic self-concept share a large proportion of common variance that may be attributable to general academic self-concept. Thus, the nested Marsh/Shavelson model of academic self-concept was proposed, and this is the most recent structural model of academic self-concept. The nested Marsh/Shavelson model integrates the assumption of the hierarchical structure of self-concept by Shavelson et al., that is, that one general academic self-concept factor is at the apex, with the multifaceted nature of academic self-concept. The nested Marsh/Shavelson model integrates the general self-concept assessed with a unidimensional scale in the form of an (incomplete) bifactor or nested model. The term nested was chosen because subject-specific self-concepts factors are nested under a general academic self-concept factor (see Brunner et al., 2010), representing general academic self-concept assessed with a unidimensional scale and the shared part of variance in the subject-specific self-concepts factors that can be attributed to a common part. Thus, the general academic self-concept is at the apex of the hierarchy. In the correlation of the subject-specific self-concepts, the shape of individuals' profile of strength and weaknesses is depicted (i.e., people think of being a more verbal or math person).

In this respect please note that there are two possibilities to translate the assumption of a hierarchical construct in a structural model in factor analysis: higher-order factor models or bifactor models.

The nested Marsh/Shavelson model utilizes incomplete bifactor models to do this whereas the Marsh/Shavelson model utilizes a higher-order factor model. The general academic self-concept assessed with a unidimensional scale is placed at the apex in the nested Marsh/Shavelson model whereas it is a subordinate in the Marsh/Shavelson model influenced by the two higher factors math domain general and verbal domain general self-concepts.

The advantage of the nested Marsh/Shavelson model compared to the Marsh/Shavelson model is that it allows researchers to *depict* and *investigate* the shape of individuals' profile formation of subject-specific academic self-concepts independent of the general academic self-concept level.

Several studies also support this model by testing it against the Marsh/Shavelson model in secondary school students with six core subjects (Brunner et al., 2010), and a restricted version with math and native language and general academic self-concept (Brunner et al., 2008; Brunner et al., 2009; Esnaola et al., 2018) showed a good fit in samples of secondary school students. However, no study investigated this model in elementary school students. To close this gap, the most recent structural model of academic self-concept, that is, the nested Marsh/Shavelson model (see Figure 2b) was tested in Study 2 in elementary school students (grades 1 to 4). More precisely, we tested the restricted version of the nested Marsh/Shavelson model with math, reading, writing, and general academic self-concepts. The nested Marsh/Shavelson model provided a good fit in our confirmatory factor analysis in Study 2. Thus, we could confirm that the model fits the data of not only secondary school students but also elementary school students. Of note, because no further school subjects were included in the nested Marsh/Shavelson model it could not be compared with the Marsh/Shavelson model with respect to a better fit in factor analysis.

Self-concept changes in structure with development (Harter, 2012a; Shavelson et al., 1976). In previous research this topic was not investigated within the nested Marsh/Shavelson model although this is better suited to depict the individuals' profile of strength and weaknesses independent from an overall level of self-perceived academic ability. Therefore, it was an aim of Study 2 to investigate structural change in the nested Marsh/Shavelson model within elementary school time. Results confirmed the results of previous research (see, e.g., Marsh, 1989a) that academic self-concept became more differentiated within the first three years of elementary school time. Note that in Study 2 structural change was shown within a theoretically better suited model, that is, the nested Marsh/Shavelson model.

To abstract from that finding within the nested Marsh/Shavelson model regarding its suitability in elementary school students and the investigation of structural change, individuals' profile of strength and weaknesses in the academic section of the self-concept, also strength and weaknesses concerning the entire self-concept can be depicted an incomplete bifactor model in the sense of a nested Marsh/Shavelson of the entire self-concept. Individuals may have, for instance, strengths and weaknesses in the academic and social domains within their self-concept profile as is the case, for example, within the academic section of self-concept regarding the strengths and weaknesses in math and native language. Under the presumption that self-esteem is empirically not distinguishable from the general self-concept higher-order factor in the Shavelson et al. model—as indicated in several studies (e.g., Craven & Marsh, 2008; Marsh & Hattie, 1996)—one could also model the entire self-concept as a nested factor model (i.e., incomplete bifactor model). This would theoretically be in line with the assumption that people have strengths and weaknesses in different self-concept areas (like academic and nonacademic) independent of their overall level of feelings toward themselves (i.e., self-esteem). Thus, individuals could see a strength in the social domain *relative* to the academic domain which is depicted in the relations between domain-specific self-concept factors in the nested factor model.

Along the same lines, individuals' self-concept profiles of strengths and weaknesses independent from an overall level as depicted in a nested model is theoretically plausible to investigate differentiation and integration regarding the entire self-concept, that is, multiple self-concepts facets not only in the academic section but also other nonacademic self-concepts like physical self-concept. Note that, in previous research for such purpose, a first-order correlated factor model or manifest scale scores were often used (see e.g., Marsh & Ayotte, 2003). These models do not allow researchers to depict an individuals' shape of the profile independent from an overall level.

As a possible reason for structural changes of self-concepts, we focused on the influence of dimensional comparisons (i.e., intraindividual comparisons within an internal frame of reference in which an individual compares different attribute of themselves with others of themselves). Dimensional comparisons as a determinant of subject-specific self-concept and academic self-concept formation was first described in the I/E model (see Möller & Köller, 2001). According to the I/E model, two comparison processes (social and dimensional) are major determinants of academic self-concept development. First, students make use of external comparisons: They compare their performance in a particular subject with the performance of other students in the same subject and with further external standards of actual achievement level. Second, students make use of dimensional comparisons, which occur when students contrast their performance in one particular school subject with their performance in other subjects. Dimensional comparisons are expected to be a source of why academic self-concept is so domain-specific, that is, math academic self-concept and native language self-concept are practically uncorrelated (see Subsection 2.4.1). Because the number of studies investigating the I/E model and dimensional effects in elementary school students is scarce compared to those investigating secondary school students (see e.g., Möller et al., 2009), we aimed to investigate the effect of dimensional comparisons and social comparisons for the development of subject-specific academic self-concepts in elementary school students (grades 3 and 4) in Studies 2 and 3. Thereby, we used different statistical approaches. In Study 3 we showed that the effect of dimensional comparison processes can be revealed using different analytic methods, that is, path analysis and the correlation comparison approach[24]. Moreover, we demonstrated in Study 2 that different structural models of academic self-concept, that is, the nested Marsh/Shavelson model and a first-order factor model, had less influence on the results regarding the I/E model. Further, in Study 3, the effects of dimensional comparison processes could be shown across samples of elementary school students and secondary school students from grade 3 to grade 8 (ages 7 to 16) and academic self-concepts assessed with different scales. Despite the finding that dimensional comparison effects are robust effects in general, less is known about moderating influences of dimensional comparison processes particularly during elementary school. To close this gap, in Study 2 we investigated if grade difference between math and native language moderates the effect of dimensional comparisons and found evidence supporting this claim. Specifically, with increasing differences between students' grades in math and native language, contrast effects of dimensional comparisons became stronger. These stronger effects may be reinforced by longitudinal negative effects of dimensional comparisons in math and the native language as posited in the reciprocal I/E model (e.g., Möller, Zimmermann, & Köller, 2014). In the reciprocal I/E model it is predicted that, within domains, achievement in math and achievement in the native language and the respective subject-specific academic self-concept have positive effects on subsequent achievement in math and achievement in the native language and subject-

[24] Quasi-experimental approach (Rost et al., 2005)

specific academic self-concept within domains but have negative effects across domains (e.g., Möller, Zimmermann, & Köller, 2014). The predictions of the model could be confirmed in several studies using path-analytic methods and nonexperimental data (see Möller et al., 2011; Möller et al., 2014; Niepel et al., 2014).

However, given the findings in Studies 2 and 3 on the effects of dimensional comparison processes it should be noted that because no experimental studies were used to investigate the I/E model, the findings may also be partly or completely explained by other variables like, for example, subjective beliefs, specifically, beliefs in a negative interdependence between math and verbal abilities (Möller, Streblow, & Pohlmann, 2006). Further, beliefs in a negative interdependence between math and verbal abilities offered and transported by teachers and parents may contribute to this effect. Parental beliefs are influence children' self-perceptions (Pesu, Aunola, Viljaranta, & Nurmi, 2016). Moreover, van Zanden et al. (2017) found support for I/E model for both students (N = 486; aged 11 to 17) and their parents. Because parents influence self-perceptions (e.g., Pesu et al., 2016) the effect of dimensional comparison may spill over to childrens' self-perceptions. Further, because achievement feedback from teachers as well as parents are sources of self-concept (Burnett, 2003; Burnett, 2004; Felson, 1989)

6.1.2.2 Stability and change in the development of self-concept facets of students in childhood and adolescence

It was an explicit aim to investigate self-concept facets on a group level and an individual level given the lack of previous research. Note, the individual level was comparably seldom investigated in previous research (see Subsection 2.2.1). Further, regarding the social self-concept of assertion, investigations on the stability and change on the individual level are totally lacking.

Therefore, this issue was addressed and investigated in Study 1 on a group level and individual level in self-esteem, academic self-concept, and social self-concept of acceptance and assertion in secondary school students from grades 5 to 8, comprising the developmental period from late childhood to adolescence. In general, there are similarities and differences in stability and change in the development of the self-concept facets under study.

All self-concept facets were similarly stable from late childhood to late adolescence. This finding contradicts the assumption by Shavelson et al. (1976) and Brown (1993) that self-esteem (as the most general self-concept) should be more stable than specific self-concept facets (i.e., academic self-concept and social self-concepts). Of note, previous research could also found no evidence for this assumption (Marsh & Yeung, 1998; Shavelson & Bolus, 1982).

Besides the similarities there are also differences between self-concept facets under study. General academic self-concept declined in mean-level during the elementary school time period (Study 2) as well as during the secondary school time period (Study 1). However, the shape of mean-level change was nonlinear which implies that the decline was weaker with ongoing grade-level. During elementary school, the subject-specific self-concepts of math, reading, and writing native language as well as general academic self-concept were assessed and displayed nearly the same pattern.

Note, that Study 1 included only students from the highest track of the German secondary school system (Gymnasium). Therefore, the big-fish–little-pond effect (BFLPE) might partly explain the decline that was found. The BFLPE is a well-replicated effect of social comparisons within a group of students (e.g., class) (e.g., Marsh, 2005). The BFLPE posits that equally achieving students have higher self-concepts when they are in a lower achieving group than when they are in a higher achieving group. The explanation of the findings are supported by the results found in a study conducted by Becker and Neumann (2017). They investigated German students across the transition from elementary to secondary school (*N* =1,617) from grades 6[25] to 9 enrolled in the highest

[25] In Germany, the transition from elementary school to secondary school takes place after sixth grade in the state of Berlin in all other federal stated of Germany elementary school time lasts 4 years.

track of German school system (i.e., Gymnasium) and comprehensive schools (i.e., Integrierte Sekundarschulen). They found that only in the first years after the transition the cross-sectional BFLPE in elementary school time was stronger for students enrolled in Gymnasium compared to students attending a comprehensive school. Stated differently, after the transition, negative elementary school effects persisted for general academic self-concept. Further, from grade 6 to grade 7 general academic self-concept increased in students attending comprehensive schools but declined in students attending a Gymnasium. However, the BFLPE of secondary school time at grade 6 was equally strong for both school-types in the short term and in the long term.

An alternative or additional explanation may be derived from stage-environment fit theory (Eccles & Roeser, 2009). In stage-environment fit theory it is assumed that a mismatch between developmental needs and educational practice (e.g., forced social comparisons in classroom, grading practices) may have contributed to that finding. Stage-environment fit theory is often taken to explain a decrease in academic self-concept during early and middle adolescence (see Section 2.2).

Despite of the development of general academic self-concept in Study 1 also social self-concepts were investigated. In previous research, investigations of social self-concept almost exclusively focused on context (e.g., peers vs. parents, school vs. home; (see e.g., Byrne & Shavelson, 1996).

The results of Study 1 reveal that social self-concept of assertion showed a different shape of mean-level change compared to social self-concept of acceptance from late childhood to late adolescence. This finding indicates that different processes underlie the development of these social self-concepts. An explanation might be that there are several ways to gain social acceptance. During adolescence it becomes increasingly possible for individuals to choose their own social niches (e.g., McElhaney et al., 2008). In these chosen niches they might feel more and more accepted over time. In contrast, as expected in the social skill approaches to social competence (cf., Rose-

Krasnor, 1997), being socially competent requires unique competencies consisting of social, emotional, and cognitive skills. Once established, they may stay constant. Particularly the time after the transition from elementary to secondary school might be important to establish the perception of having the ability to assert oneself in social interactions. Individuals have to find their place in new social contexts (i.e., the classroom) where it is not only of importance to get along with unknown peers but also to assert oneself.

After discussing the findings on stability and change in academic self-concept development, the results on the causal flow between self-esteem, general academic self-concept, and social self-concepts of acceptance and assertion throughout adolescence is discussed in the following section.

6.1.2.3 Causal flow between self-concept facets

There is an ongoing debate on how general self-concepts (e.g., self-esteem as the most general self-concept) and specific self-concept facets are causally related over time (see e.g., Trautwein, 2003). Does changes in one self-concept facet determines change in another self-concept facet? Note, for fostering positive and at the same time realistic self-esteem, it is of significance to find out determinants of self-esteem which are theoretically and empirically based (see Harter, 2012b). Given the hierarchical structure of self-concept posited in the Shavelson et al. model, the effects of causal flow between self-esteem and specific self-concept facets (i.e., academic and social) can be described as *bottom-up* effects if the direction of longitudinal relation goes from specific self-concepts to general self-concept (e.g., from social self-concept of acceptance to self-esteem), *top-down* effects if that causal flow mainly goes from self-esteem to specific self-concepts, and *reciprocal* if the direction of causal flow is bottom-up and top-down at the same time.

Theoretical assumptions about the causal flow between general and specific self-concept facets are somewhat contradictive and not consistent (see Marsh & Yeung, 1998). To date, there are only few longitudinal studies investigating this question, and their findings are inconsistent (see e.g., Marsh & Yeung, 1998; Trautwein, 2003).

Therefore, it was an explicit aim in Study 1 to investigate the causal interplay between self-concept facets comprising self-esteem, general academic self-concept, and social self-concepts of acceptance and assertion across a long time span and multiple waves (i.e., 6 time points) comprising late childhood to late adolescence in secondary school students (grades 5 to 11). This extends previous research that has used longitudinal data comprising only two measurement waves (see Marsh & Yeung, 1998; Shavelson & Bolus, 1982; Wagner & Valtin, 2004) or three measurement waves within one year (see Trautwein, 2003).

Further, the causal interplay between specific self-concepts is rarely investigated and discussed. Trautwein (2003) designated the effects between specific self-concept facets transdimensional. These effects are frequently investigated with respect to the influence of dimensional comparison processes (i.e., intraindividual comparisons of attributes of a domain with attributes of another domain) in the academic section of the self-concept (see the previous Subsection). With dimensional comparison theory (see Möller & Marsh, 2013), the effects of dimensional comparison processes were also discussed within other domains like the academic domain and the social domain. Möller and Husemann (2006) found that individuals conduct dimensional comparisons in different areas of life and also between the social and academic domains. Thus, transdimensional effects can also be considered as, inter alia, being a result of dimensional comparison processes.

From Study 1 it can be concluded that there is a causal flow between self-concept facets, to a small degree, from late childhood to late adolescence in students from the highest track of the German secondary school system. We found a relatively balanced pattern of bottom-up and top-down effects as well as reciprocal effects but only to a small degree. Thus, neither a predominance of bottom-up effects nor top-town effects could be shown. In Study 1, former social self-concept of assertion positively predicted later self-esteem at a later time point between two time lags. Regarding social acceptance, no bottom-up effect was found. Note that social self-concept of assertion was less targeted in previous research. Further, no bottom-up effects were found for social self-concept of acceptance. This is somewhat in contrast to a recent study: Tetzner et al. (2016) investigated the causal interplay between self-esteem, social acceptance, and achievement in a sample German secondary school students (*N* = 7,977; 3 time points; start of grade 7, end of grade 7, end of grade 10). They found bottom-up effects from the end of grade 7 to the end of grade 10. However, this study differed in several ways from Study 1. Tetzner et al. controlled for achievement, investigated a different time span, their sample stems from different tracks, and finally, they did not include social self-concept of assertion in the analysis.

Further, in Study 1 some evidence for transdimensional effects between academic self-concept and social self-concept of acceptance and assertion were found. Note that these effects are partly replicate the results from a previous study by Preckel et al. (2013) who investigated only the causal interplay between social self-concepts and academic self-concept with less time points in an partly overlapping sample used in Study 1. Results that go beyond that previous study indicated positive effects from students' earlier social self-concept of assertion to their later social self-concept of acceptance from the beginning of grade 5 to 11 weeks later. Moreover, there was a negative effect from earlier general academic self-concept to later social self-concept of assertion from grade 8 to grade 11 and a negative effect from earlier social self-concept of acceptance to later general academic self-concept from grade 6 to grade 8.

To conclude going beyond previous research we are able to draw conclusions about the pattern of causal flow between self-concept facets from late childhood to late adolescence which could hardly be drawn from previous studies because no previous study covered this time totally. To this end, the findings from Study 1 contributes to existing knowledge. To conclude, going beyond previous research we are able to draw conclusions about the pattern of causal flow between self-concept facets from late childhood to late adolescence which could hardly be drawn from previous studies because no previous study covered this time frame entirely. Thus, the findings from Study 1 contribute to the extant knowledge.

However, also in the present study as well as in previous research, results diverge between studies and it was not possible to resolve the inconsistencies in previous research. A critical aspect in this respect is the chosen analytical method in the previous studies investigating self-esteem, academic self-concept, and social self-concept to investigate the causal flow, namely cross-lagged models. Cross-lagged panel models are justified to a certain degree to allow conclusions about the direction of causal flow between variables. But, no theory of change is implied in these models (Selig & Little, 2012). This can lead to problems when a theoretical assumption is specific to a theory of change.

Hamaker et al. (2015) criticized cross-lagged panel effects in that cross-lagged and autoregressive effects can be affected if different trait-like stabilities in constructs under investigation exists. In Study 1, we found different trait-like stability in self-concept facets that may have biased the results of cross-lagged effects of interest. This might also be the case in previous studies. Another aspect is that the effects between the studies cannot be easily compared due to different time lags.

Besides these critical aspects of possibly biased effects due to trait-like stabilities of diverse self-concept facets under investigation and the difficulties to compare the results from different studies due to different time lags, when making theoretical predictions and testing these with statistical methods regarding the causal flow between self-concept facets it might be important to *differentiate* between the trait and state component of self-concept facets in general. Stated differently, if the causal flow should concern the trait or state component of self-concept facets is theoretically of importance for predictions for the causal flow (see Parker, Marsh, Morin, Seaton, & van Zanden, 2015).

An important hint in this respect is given by a study by Parker et al. (2015). They investigated the development of math and native language change and tested the implicit assumption of the dimensional comparison theory that changes in math should negatively predict changes in native language and vice versa in N = 2,781 Australian high school students from grade 7 to grade 11 (cohort-sequential design with 10 time points; 6 months apart). They used three different analytical methods: (1) autoregressive cross-lagged models (i.e., cross-lagged panel models) to test the temporal ordering of the associations between native language self-concept and math self-concept; (2) latent growth curve models to test hypotheses about the direction of growth in these two domains and the relationship between their growth trajectories; and (3) autoregressive latent trajectory models to explore whether the relationship between native language and math self-concept differs at state residual (fluctuations from trajectory) and trait levels. They found a significant longitudinal negative relation between math and native language self-concepts over time, but only for the state residual components of self-concepts (i.e., in the autoregressive latent trajectory models) and not in the cross-lagged panel models. This finding has theoretical as well as methodological implications. Theoretically, one would expect only a causal relation between self-concept facets in the state component but not in the trait component[26].

Parker et al. (2015) concluded:

> “These results for the state residual components of self-concept may suggest that dimensional comparisons relating to fluctuations from trait trajectories in self-concept may act as

[26] Note that the state and trait distinction is here statistically, thus, method-driven. State- and trait-related variance components can be estimated in statistical measurement models based on repeated measurements of observed variables. This has to be distinguished from questionnaires and measures which explicitly aim to assess state (i.e., “At the moment, I feel like I am worthless”) and/or trait (i.e., “I am a worthless person”) of a given individual.

> a mechanism to return overall academic self-concept back to a stable growth after events (e.g., doing much better than expected on a math test, or failing an English test that you believed you would do well on) create changes in self-perceptions. Put simply, an event that leads to a student doubting their ability in math may result in a rise in their English self-concept, thus helping to promote a stable growth trajectories in academic self-concept. In this way, the extreme multidimensionality of academic self-concept and dimensional comparisons may help act as resilience mechanisms" (p. 187).

Thus, in line with the explanation offered by Parker et al. to interpret the finding with respect to the relations between math and English self-concept to general academic self-concept. The train of thought might also apply to social self-concept and academic self-concept with respect to self-esteem. In this context the assumptions made in the compensatory model of self-concept by Marx and Winne (1980) (see Byrne, 1984) should be noted. In the compensatory model (Marx & Winne, 1980; Winne & Marx, 1981) it is assumed that a low academic self-concept could be compensated by a high social self-concept and vice versa with respect to achieve and maintain a high level of general self-concept (i.e., self-esteem). Indeed, Möller and Husemann (2006) showed in two diary studies with 67 university students (Study 1) and 65 high school students (Study 2) investigating dimensional comparisons in everyday life that individuals use such contrasting comparisons between social and academic areas to improve mood and global feelings of worth (i.e., self-esteem). To bring together the explanation of findings by Parker et al., the finding from Möller and Husemann (2006), and the theoretical model by Marx and Winne (1980), it seems reasonable to conclude that contrasting effects of dimensional comparisons between social and academic self-concept may be seen as a resilience mechanism or self-regulating mechanism to maintain a stable growth trajectory in self-esteem.

In conformity with this conclusion seems to be the theoretical assumption proposed by Hattie (2004) regarding the relation between self-concept facets should be quoted. Hattie considered the interplay between self-concept facets from the view of strategies and mechanism (i.e., processes) to maintain self-concept. In his view, self-serving strategies (e.g., self-enhancing, self-handicapping, self-verification) that related self-concept facets with each other are more important when questioning how self-concept facets are causally related across time.

Extracting from Parker et al.'s (2015) findings, the interplay between self-concept facets across time may be more complex than theoretically expected and the distinction between state and trait component in self-concept facets and the chosen analytical methods may play a role in interpreting findings. In this respect, further research is needed with methods allowing the separate investigation of state and trait components of self-concept facets to shed light on the pattern of relation between self-esteem and specific self-concept facets as well as between specific self-concept facets. Further, expanding on Hattie's (2004) theoretical view, a process perspective from self-strategies (e.g., self-enhancing, self-handicapping, self-verification) to maintain self-esteem may contribute to a better understanding of how and if self-concept facets are longitudinally related in the state and trait component of self-concept facets.

6.2 Implications

6.2.1 Theoretical implications

There were four major theoretical implications derived from the reflection of central findings. The first theoretical implication concerns the use of incomplete bifactor factor models to depict the structure of self-concept. It was argued that in the nested Marsh/Shavelson model of academic self-concept (nested factor model; incomplete bifactor model), hierarchy and multidimensionality is combined in a theoretically very plausible way (see Section 2.1 and Subsection 6.1.2.1). That is, one could have strengths and weaknesses in self-concept areas independent from an overall level in general self-concept. Thus, nested factor models like the nested Marsh/Shavelson model may contribute to the translation of plausible theoretical assumptions regarding the definition of a general self-concept and the hierarchical aspect of self-concept in a factor model representing the structure of self-concepts—in the academic section and the entire self-concept. In particular, from a theoretical perspective for research on structural changes (i.e., differentiation and integration) in self-concept development, it is a suitable model because a nested factor model depicts individuals' profiles of strength and weaknesses in self-concept. Thus, future research should use this model for these purposes.

As outlined in the previous section, because of the limited number of school subjects examined during elementary school, it was not possible to directly compare the nested Marsh/Shavelson model of academic self-concept with the Marsh/Shavelson model of academic self-concept regarding model fit in confirmatory factor analysis (CFA). Thus, future studies should test the nested Marsh/Shavelson model of academic self-concept with further subjects taught in elementary school. This would further support the structural assumptions made by the nested Marsh/Shavelson model. In Germany, additional elementary school subjects may include foreign language, religion, music and art, physical education and social studies, writing and reading. The pattern of correlation between subject-specific self-concept factors should correspond to the assumption of the dimensional comparison theory that similar (near) subject-specific self-concepts should correlate higher than more dissimilar (far) one (see Figure 2).

Second, experimental research is needed to verify that dimensional comparisons are the cause for a multidimensional academic self-concept in elementary school students and secondary school students found in Studies 2 and 3.

Third, considering the results of Study 1 leads to the conclusion that social self-concept should be investigated regarding its multidimensionality within content (i.e., assertion and acceptance). Most previous studies focused on social self-concept acceptance and neglect self-perceptions of one's social skills in social interactions. This may be explained at least partially because many self-concept measures predominantly assess social acceptance, popularity, and the quality of social relations (Berndt & Burgy, 1996; Preckel et al., 2013) or other aspects of social competence. Frequently implemented multidimensional self-concept questionnaires (see Donnellan, Trzesniewski, & Robins, 2015) include the series of the Harter Self-Perception Profiles (see Harter, 2012b) and the series of Self-Description Questionnaires (e.g., Marsh, 1990a). These measurement instruments do not assess social self-concepts separately within the content. In the Harter Self-Perception Profiles from 1988 (for children and adolescence), only social self-concept of acceptance is assessed. In the revised version from 2012 (see Harter, 2012b), Harter acknowledged the aspect of social competence and incorporated it in the series of Self-Perception Profiles (childhood to adolescence). Harter (2012a) described the items of the scale social competence as items that "refer to knowing how to make friends, having the skills to get others to like oneself, knowing what to do to have others accept you, understanding what it takes to become popular" (p. 3). Thus, the items to assess social self-concept now include additional items to capture social acceptance as well as aspects of social competence and popularity. Therefore, the scale was renamed social competence (formerly: social acceptance). Note that in the SDQ-III (Marsh, 1990d), the scale assessing social self-concepts also comprise items regarding social acceptance as well as aspects of social competence (how to make friends) and popularity, too. However, within both measurement instruments, these are not seen as separate aspects of social self-concept. Moreover, the items that Harter and Marsh used (e.g., "how to make friends") to assess social competence perceptions may be a separate (but probably related) aspect from those that were used by Fend and Prester (1986) (e.g., "Sometimes I don't say anything although I am right"). The social skill *assertion* related to a kind of communication that aims to *prevent the loss* of reinforcement rather than gaining success in social interactions (Nangle et al., 2009). This is contrary to “how to make friends,” which might relate to the social skills that aim to build relationships and therefore aim to enhance the reinforcements by others. Thus, it is doubtful that with items to assess ”knowing how to make friends, having the skills to get others to like oneself, knowing what to do to have others accept you, understanding what it takes to become popular” (Harter, 2012a, pp. 3-4) the construct of self-concept of social competence is in an appropriate manner captured. To conclude, social self-concept should be further investigated as a multidimensional construct within content. Thereby, social self-concept of assertion should be included in measurement instruments that aim to assess social self-concept of competence.

Third, the results from Study 2 indicated that multidimensionality within the native language self-concept should be considered. Reading and writing self-concept are two separate facets of native language academic self-concept in elementary school. Further, reading and writing self-concept showed different effects within the I/E model. Thus, it seems to be important to include both aspects in future studies that aim to investigate dimensional comparison effects within the I/E model.

Fourth, regarding the causal flow between self-concept facets it might be important to include state and trait components of self-concept in the theoretical assumptions on the causal flow and to use statistical methods that are able to disentangle these effects[27].

Further, to gain a better comparability of effects between studies investigating the causal flow with different time lags methods should be used that account for this point. Statistical methods that model the time in a continuous way seems to be particular promising because the effect size of change can be calculated for every arbitrary time-span allowing to compare effects between studies

[27] Further, it is important to use diagnostic measurements (e.g., a questionnaire) that are able to assess states and traits.

using different time lags. Further, within continuous time modeling trait-like influences can also be handled (see Delsing & Oud, 2012; Oud & Voelkle, 2014; Voelkle, Oud, Davidov, & Schmidt, 2012). Apart from that, the influence of self-strategies in the complex interplay between self-concept facets should be considered more clearly in future research.

6.2.2 Practical implications

The structure and development of self-concept and its determinants are prerequisite for understanding self-concepts. Therefore, this knowledge is extremely important to plan and implement interventions to promote positive self-concepts in school-age children and adolescents.

The findings of Studies 2 and 3 highlight that social and dimensional comparisons are determinants of (academic) self-concepts that should be considered by teachers, parents, educational counselors, and in the planning and implementation of interventions to foster positive self-concepts.

However, dimensional comparisons have been referred to as a double-edged sword (see Marsh & Möller, 2013) that lowers a student's self-concept in the weaker domain while raising it in the stronger domain. That is, a student's comparison of his or her achievement in math with his or her own lower achievement in the native language (downward dimensional comparison) may enhance his math academic self-concept, while comparing his or her achievement in math with a higher achievement in the native language (upward dimensional comparison) may lower his or her math academic self-concept. Thus, every upward comparison from target to standard is accompanied by a downward comparison from standard to target. Möller and Marsh (2013) stated that: "even negative effects of achievement in a better off domain on self-concept in a weaker domain seem reasonable, as they prevent wrong decisions" (p. 550). Thus, the effects of dimensional comparison decisions may prevent false decisions to be made in educational choices and, therefore, the academic career. However, this may only hold true under the assumption that a person has a profile of strengths and weaknesses that correspond to his/her “real” strengths and weaknesses. For this, it is important that a student receives accurate achievement feedback that relies on the objective performance. Feedback from significant others (e.g., teacher, parent, peers) and reflected appraisal are seen as sources of self-concepts (Harter, 2012b; Marsh, 1986; see Shavelson et al., 1976). Regarding the formation of academic self-concepts, direct or indirect achievement feedback from teachers is a source of academic self-concepts. Therefore, teachers should be aware of the effect of dimensional comparisons and rely on the objective performance in a specific subject when giving feedback that is as accurate as possible.

Moreover, comprehensive achievement-related feedback has been found to be effective in fostering academic self-concepts as one meta-analysis on 145 studies on interventions to enhance self-concept has demonstrated (O'Mara et al., 2006).

However, despite the possible positive effects of making good decisions, the effect of dimensional comparisons regarding achievement development in the specific subjects seems unwanted. There is remarkable evidence from previous research highlighting that subject-specific academic self-concept predicts achievement and vice versa (reciprocal effects model, REM; e.g., Marsh & Martin, 2011; for a meta-analysis see Huang, 2011). Further, there is also evidence for negative longitudinal effects of dimensional comparisons in math and native language self-concepts as posited in the reciprocal I/E model (reciprocal I/E model; Marsh & Craven, 2006) (see Subsection 6.1.2.1). Thus, negative side-effects of dimensional comparisons between math and native language can lead to a poorer scholastic performance. Such effects seem to be unintended side-effects because a high achievement in more or less all school subjects is similarly desired to obtain an average good grades and, in the end, a good graduation which permits individuals to take part in higher and university-level education. Of course, students can choose particular subjects to set priorities (e.g., the choice of advanced courses in upper secondary school), but core subjects like math and native language are mandatory subjects.

To conclude, it should be kept in mind that, within the academic self-concept section, contrast effects of dimensional comparison processes (that was found in Studies 2 and 3) might limit positive side effects in elementary school as well as in secondary school when fostering subject-specific self-concepts. Because there are several possibilities to foster academic self-concepts (see cf., O'Mara et al., 2006), it seems a good advice to circumvent the topic of dimensional comparisons. Therefore, a good choice might be to use indirect approaches that do not explicitly target to enhance achievement and academic self-concept in a single school subject (see Trautwein & Möller, 2016). For example, a possible intervention might be attribution (re)training (e.g., Siegel Robertson, 2000) in which students were taught adaptive styles of attributions after success and failure. Central objectives of attribution (re)training are to (a) provide students with realistic knowledge of their own abilities, and (b) provide and/or altering, in a positive fashion, their competence related self-perceptions and motivational performance behaviors (see Siegel Robertson, 2000; Sukariyah & Assaad, 2015).

Further, because the development of self-concept facets (academic and social self-concepts) in late childhood to late adolescence was investigated on the group level and individual level in Study 1, a suggestion can be made (with limitations) as to whether interventions should focus on all or only some individuals. This is guided by whether a possible negative change in a self-concept facet concerns the majority of individuals or only some of them. In this regard note that only the trajectory (level change) was considered. Of note, initial and overall level of the self-concept facets may be unrealistically low or high or based on undesirable sources. This aspect, however, was not considered in Study 1 because no external criteria for academic self-concept (e.g., achievement measures), social self-concepts (e.g., social competence evaluated by peers or teachers), or adverse sources of self-esteem (e.g., narcissism) were included in the analysis. Noteworthy, with respect to academic self-concept, previous research indicated that even an overestimation of academic ability can be favorable for the positive development of achievement (Praetorius, Kastens, Hartig, & Lipowsky, 2016; but see Trautwein & Möller, 2016).

Thus, whereas nearly all individuals showed a decline in general academic self-concept throughout secondary school, intervention programs to foster academic self-concept should target all individuals. It can be concluded from the result of Studies 1 and 2 that general academic self-concept declined during elementary school and secondary school. Therefore, interventions fostering academic self-concept should be implemented at the beginning of elementary school. Because evidence exists from previous research that academic self-concepts and achievement are reciprocally related (reciprocal effects model, REM; e.g., Marsh & Martin, 2011) across different age ranges (see the findings from the meta-analysis by Huang, 2011), and this concerns students of different achievement levels, that is, low, medium, and high achieving students (Susperreguy, Davis-Kean, Duckworth, & Chen, 2017), a decline in mean-level in academic self-concept may be seen as detrimental with respect to achievement development. This is an unwanted outcome. Thus, academic self-concept should be fostered in students. Because the sample of Study 1 consisted of students from the highest track of secondary school system in Germany, the BFLPE might have contributed to a decline after transition until middle adolescence (see Subsection 6.1.2). In the highest track there are more possibilities for social upward comparisons, which might impact academic self-concept. Social comparisons can have positive or negative effects on motivation and achievement. The comparison with a better of student can have positive or negative consequences. The negative effect of social comparisons can be buffered if the feedback of teachers was accompanied with motivational cues (you can make it if you put in effort) or with informational values (the student used an effective learning strategy that is generally helpful) (see Trautwein & Möller, 2016).

Further, regarding social self-concept of acceptance and assertion, it can be concluded from the results of Study 1 that only a minority of individuals showed a decline from late childhood to adolescence. Here, individualized interventions are appropriate. However, the causes for a decrease in social self-concept of acceptance can vary widely from a lack of social skills in general or in specific situations (e.g., low assertion in social interactions) to being a victim of bullying. If the cause is a lack of social skills, social skill training programs may be appropriate (see for an overview of interventions: Nangle et al., 2009). These programs may also be a good choice to foster the social self-concept of assertion.

In Study 3 evidence for the I/E model with an alternative approach using cross-sectional and non-experimental data proposed by Rost et al. (2005; see also Schilling et al., 2004), that is, the correlation comparison approach (CCA) was found. Application of the CCA across elementary school and secondary school students (grades 3 to 8) showed that the effects of dimensional comparisons are stronger when achievement differs between subjects. Using this approach is in line with the principle of a minimally sufficient analysis that states that “if the assumptions and strength of a simpler method are reasonable for your data and research problem, use it” (Wilkinson, L. & Task Force on Statistical Inference, American Psychological Association, Science Directorate, 1999, p. 598). Compared to path analysis (which is most often used to investigate the I/E model with cross-sectional and nonexperimental data), the CCA is less prone to multicollinearity and, if used with manifest scale scores, is easy to implement and operate on available data. Using the CCA, our study showed that the effects of dimensional comparisons are stronger when achievement differs between subjects. Thus, this approach complements I/E model research with cross-sectional and nonexperimental data.

6.3 Strengths and limitations

A major strength of the present dissertation is the latent modeling of self-concept facets[28] in all studies. In doing so, measurement error and method effects were controlled for (see e.g., Kline, 2016). Further, we tested specific forms of measurement invariance that are prerequisite for testing hypotheses on latent means, latent covariances, and correlations. Note that nearly only previous self-concept research tested for measurement invariance (see Marsh & Ayotte, 2003) before investigating structural changes (i.e., change in the correlations between self-concept facets) in self-concept development. Moreover, in Study 1 it was drawn on recent developments in confirmatory factor analysis to identify and scale latent factors by using the effect-coding method (Little, Sleggers, & Card, 2006). The effect-coding method is advantageous for two reasons: First, the variance estimation of the latent factor is equivalent to the average of the items reliable variance. Second, the latent mean of a construct at a single time point can be interpreted as the average of the items mean weighted by the factor loadings (Little, Preacher, Selig, & Card, 2007). Thus, it is a more accurate estimate of the “true” mean value than an arbitrarily chosen indicator mean (as in the marker variable approach) (Little et al., 2006).

Despite this, a further strength is the data basis for the three studies. This concerns (a) the sample sizes and the (b) quasi-representativeness of German students. Specifically, for the investigation of individuals’ academic self-concepts from early to late childhood, we used data from a large quasi-representative national data set comprising five federal states of German elementary school students from grades 1 to 4 (Total N = 3,779). Further, to investigate multiple self-concept facets from late childhood to late adolescence, we used large data sets from two studies comprising German secondary students from four federal states.

[28] Study 3 was investigated within a manifest as well as a latent variable framework.

Besides the strength of the data basis of the three studies, there are also limitations. The samples used in the three studies are partly overlapping with respect to the three projects (see Table 1). Further, all samples comprise students attending schools in one country, that is, Germany, and this has several implications for the generalizability of the results with respect to other countries. First, regarding the effects of dimensional comparisons investigated in Studies 2 and 3 about the same effects can be expected regarding samples stemming from other countries because there is evidence from previous research that the I/E model and therefore the effect of dimensional comparisons and social comparisons generalizes across different countries and cultures (Marsh & Hau, 2004; Möller et al., 2009).

Nonetheless, educational systems differ between countries and within countries in length of schooling, transition times, and curriculum (European Commission/EACEA/Eurydice, 2017). Contextual factors (i.e., school transition) can be expected to explain interindividual differences in intraindividual change, and mean-level change in diverse self-concept facets (e.g., Schaffhuser et al., 2016).

Second, there are limitations regarding the findings in students in late childhood to late adolescence enrolled in secondary education schools. Only students from the highest track of the German secondary school system were investigated. Tracks differ in several ways (e.g., curriculum, level of demands, contexts, cultures), and that might have an influence regarding findings of Study 1 particularly with respect to the decline in general academic self-concept (see the reflection of the results in Study 1 in Subsection 6.1.2).

Moreover, there are strengths and limitations of the study design in all three studies. Limitations concern the nonexperimental data used in all three studies and the cross-sectional design for investigating structural change and change in mean-level in in elementary school students in Study 2. Note that it is not unusual to use cross-sectional data to investigate developmental issues (see

Marsh & Ayotte, 2003). However, there are restrictions regarding the internal and external validity; confounds with cohort effects and history-related effects could not precluded (see e.g., Robinson et al., 2005). Moreover, with cross-sectional data it is not possible to investigate individual-level change, that is, interindividual differences in intraindividual change or ipsative stability (Robinson et al., 2005)

Further, we used cross-sectional and nonexperimental data to investigate dimensional comparison processes in Studies 2 and 3. However, this is in line with the majority of previous research on dimensional comparison processes posited in the I/E model (see Möller et al., 2009). Thus, it could only be assumed that children used dimensional comparisons and social comparisons and that dimensional comparisons contribute to a differentiation (stronger dimensionality) during elementary school. To draw this conclusion, experimental study designs are needed. Although there are some studies (cf., Marsh et al., 2017) that investigated the effects of dimensional comparison processes in experimental studies, and all of these investigated individuals beyond elementary school. However, it should be noted that in Studies 2 and 3, dimensional comparisons as formulated in the I/E model were investigated in a seldom investigated age range, that is, in elementary school students, compared to older age groups (i.e., secondary school students).

A further limitation due to nonexperimental data concerns Study 1 in which the direction of change between self-concept facets was investigated with longitudinal data. Effects regarding the direction of change must not definitively imply causality (see e.g., Newsom, 2015). Of note, the experimental manipulation of self-concept facets in individuals seems difficult or even impossible. Thus, the present study cannot establish causality between self-concept facets, but the temporal nature of the data does give some support for the causal timing of the effects (see also Marsh & Yeung, 1998). Therefore, further studies are needed to replicate the findings of Study 1 with other samples. Moreover, the time lags in Study 1 are not equally spaced, which is problematic when comparing effects (cross-lagged and autoregressive effects) between time points within the study.

Besides the strengths and limitations regarding the samples and the study designs of the three studies, a strength can be seen in the inclusion of social self-concept of assertion in the investigation of social self-concept. Specifically, social self-concepts were investigated separately within the content (acceptance vs. assertion) in Study 1, which was seldom the case in previous research. However, it should be noted that in the present dissertation multidimensionality regarding the context (e.g., school, home) (see Byrne & Shavelson, 1996) and the division in "opposite-sex" and "same-sex" was neglected in Study 1 (see Subsection 2.1.3). Content specific social self-concepts may differ between contexts and sex, too. It is possible that an individual can have the perception of being assertive with parents (i.e., at home) but not so assertive with peers (e.g., at school).

In several ways further limitations concern the spectrum of self-concept facets under study. In the present dissertation, self-esteem, general academic self-concept, math and verbal academic self-concept (in elementary school students: reading and writing), and social self-concept of acceptance and assertion were investigated. However, other nonacademic self-concept facets, (e.g., physical appearance) and additional self-concept facets regarding the Shavelson et al. model and its refinements (see Subsection 2.1) were not included.

Moreover, in elementary school student samples in Study 2 and 3, native language self-concept was comprised of writing self-concept and reading self-concept. Previous research often used only reading self-concept to represent native language self-concept. Because previous research indicated that reading did not sufficiently cover self-concept in native language (Arens, Yeung, & Hasselhorn, 2013) and reading and writing are core skills during elementary school (see Ehm et al., 2014), we included in the studies reading as well as writing academic self-concept. However, we did not capture the entire multidimensional native language self-concept facets which comprise listening and speaking, too (see Arens & Jansen, 2016). Moreover, because school subjects other than math and native language (i.e., reading and writing) were not included in the study, only a restricted version of the nested Marsh/Shavelson model of academic self-concept could be tested. Therefore, it was not possible to compare the model fit of the nested Marsh/Shavelson model with that of the Marsh Shavelson model, which includes additional school subjects like social studies and physical education (see Figure 2).

Furthermore, the effects of dimensional comparisons outlined in the *dimensional comparison theory* (Marsh et al., 2015; Möller & Marsh, 2013) of near and far comparisons between different school subjects in elementary school students as well as in secondary school students could not be investigated because in all studies investigating dimensional comparison effects, only math and native language self-concept (reading, writing) and no further subjects were investigated. This also concerns the moderating influence of the grade difference between subjects.

An overall strength of the present dissertation is that dimensional comparison processes as determinants of academic self-concepts were intensively investigated. In particular, Study 2 focused on a moderator for dimensional comparisons with the I/E model. With this in mind, it is complied with the request (see Möller & Marsh, 2013) to further clarify which factors moderate dimensional comparisons contributing to the theoretical rationale of dimensional comparison theory (Möller & Marsh, 2013). Thus, in Study 2 it was shown for the first time that grade difference acts as a moderator for dimensional comparisons in students in grades 3 and 4, that is, the higher the grade difference, the stronger the effects of dimensional comparisons were.

Finally, a major limitation of the present dissertation is that, in Study 1, no outcome variables and predictors regarding group-level change (mean-level change and shape of the change) and individual-level change (i.e., interindividual differences in intraindividual change) in self-esteem, general academic self-concept, and social self-concept of acceptance and assertion were included.

So, after having outlined the strength and limitations of the studies in the present dissertation in the following an outlook regarding future self-concept research is given based on and going beyond the limitations.

To summarize, major strengths of the present dissertation include the latent modeling framework, the sample, the study design, the inclusion of seldom investigated self-concept facets, and the investigation of change and stability on sample-level change and individual-level. Limitations of the present studies include the number of self-concepts investigated, the disregard of outcome variables and predictors of group-level change and individual-level change.

6.4 Outlook

Theoretical and practical implications of the findings from the three studies have already been outlined in Subsections 6.2.1 and 6.2.2. In the following an outlook regarding self-concept research in future based on and going beyond the limitations. Thereby, recent advances in structural equation modeling will be considered.

6.4.1 Structure of self-concept and its structural change

Moreover, within the last years, new developments in factor analytic methods offer new opportunities to model the structure of constructs within a latent variable framework. In the meantime, exploratory structural equation modeling (ESEM) was developed (Asparouhov & Muthén, 2009), which integrates Exploratory Factor Analysis (EFA) in a Structural Equation Modeling (SEM) framework, "making methodological advances typically reserved to CFA and SEM available for EFA measurement models" (Arens & Morin, 2016, p. 119). Further, bifactor Bayesian SEM and bifactor ESEM was developed (Muthén & Asparouhov, 2012). The argumentation for cross-loadings is that correlation between factors may probably be too high if not considered (see Arens

& Morin, 2016). Even when the CFA model fits well, factor correlations will be at least somewhat inflated unless all cross-loadings are close to zero (see Marsh, Liem, Martin, Morin, & Nagengast, 2011; Marsh, Nagengast et al., 2011). This is (a) due to the error-prone nature of indicators (e.g., items) used to assess self-concepts with a questionnaire, and (b) due to the hierarchical structure of a construct (i.e., self-concept). In particularly, in the presence of partially overlapping self-concept facets (e.g., foreign language, native language, peer and parent self-concepts) this may also apply to bifactor model where for a hierarchical structure of a construct is accounted for in a factor model. In two previous studies a complete bifactor ESEM and Baysian SEM to model the structure of the entire self-concept with the SDQ-I provided a good fit to the data in factor analysis (Arens & Morin, 2016; Morin, Arens, & Marsh, 2015).

In Study 2, however, the structure of academic self-concept was investigated using CFA and cross-loadings were not allowed as depicted in the original nested Marsh/Shavelson model of academic self-concept (see Figure 2b). Again, it was the aim to test the *original* nested Marsh/Shavelson model. But, given the advances in latent framework modeling to combine CFA and EFA, the nested Marsh/Shavelson model could be modeled using bifactor Bayesian SEM or bifactor ESEM. Because Bayesian SEM is advantageous over ESEM in cases of prediction or correlation to other variables (see Asparouhov, Muthén, & Morin, 2015), which is the case within the I/E model (achievement indicator predicts academic self-concepts), Bayesian SEM might be the promising approach in this respect.

A nested Marsh/Shavelson model of academic self-concept and/or the entire self-concept modified by the new statistical possibilities (ESEM and Bayesian SEM) would be advantageous in investigations of structural changes (i.e., differentiation and integration) because, to investigate structural changes, the change in the *correlations* between specific subject-specific self-concepts and/or specific self-concept facets is tested. It would be interesting to determine to which extent the kind of structural model (original vs. nested Marsh/Shavelson model) affects correlations between self-concept facets compared across time and, therefore, conclusions regarding structural change. Thus, it seems a promising avenue for future research to investigate structure as well as structural change within self-concept development using a nested Marsh/Shavelson model that is modeled with bi-factor ESEM or Bayesian SEM.

6.4.2 Comparison processes as determinants of academic self-concepts

As a major limitation the nonexperimental data in all studies of the present dissertation was mentioned in Subsection 6.1.2. To date, to the best of my knowledge, no experimental study investigated a) dimensional comparison processes in elementary school, b) the interplay between social and dimensional comparisons, and c) the interplay between social, dimensional, and temporal comparisons in elementary school students. Only one quasi-experimental study by Dickhäuser and Galfe (2004), investigating German elementary school students in grades 3 and 4 (N = 379), found no effect of dimensional comparisons or temporal comparisons with the exception of social comparisons in math self-concept. This is in contrast to the few studies that showed the effects of dimensional and social comparisons in nonexperimental and cross-sectional studies in elementary school students (see e.g., Ehm et al., 2014; Pinxten et al., 2015). However, a reason for inconsistent results could be that the chosen design in the study by Dickhäuser and Galfe (2004) contributed to this because students could freely choose the subject with which they compared their achievement in math. There are some studies with students beyond the elementary school years that use an experimental design to investigate dimensional comparison processes. For instance, Möller and colleagues conducted several experimental and quasi-experimental studies on social and dimensional comparisons in secondary school students and undergraduates (see cf., Marsh et al., 2017). For instance, Müller-Kalthoff, Jansen et al. (2017) investigated in 3 experiments and 2 quasi-experimental studies the relative effects of upward and downward dimensional comparisons as well as their net effect. In all studies German secondary school students or undergraduates were investigated. In Studies 1 ($N = 149$), 2 ($N = 150$), and 3 ($N = 300$), participants were asked to infer self-concepts of fictitious students after receiving experimentally manipulated information about their achievements in math and native language subject; in Studies 4 ($N = 2{,}268$) and 5 ($N = 20{,}662$) students assessed their own self-concepts in German and math. In all studies, downward dimensional comparisons resulted in higher self-concepts, whereas upward dimensional comparisons led to lower self-concepts. The net effect of dimensional comparisons was always found to be not statistically different from zero. They concluded that “results indicate the effect patterns to be rather universal as they were stable across different samples, domains, achievement situations, research designs, and types of assessment” (p. 1029).

Recently, Müller-Kalthoff, Helm, and Möller (2017) investigated the effects of social, dimensional, and temporal comparison and their interplay in two experimental studies (samples: university students and 11th grade German secondary school students) and one quasi-experimental study (sample: 9th and 11th grade German secondary school students). They found that all three comparison processes influenced academic self-concepts separately with social comparisons exerting relatively the strongest influence. No interaction effects between the three comparison processes on academic self-concept were obtained. However, it could only be speculated how the relative influence of the comparison processes and their interplay during elementary school is. Thus, it should be the task of future research to investigate and replicate these findings using elementary school students as samples.

To sum up, less studied to date is the interplay between these three comparison processes and the relative impact using experimental designs (see Möller, 2016; Möller & Marsh, 2013). This concerns elementary school students as well as secondary school students.

Thus far, main evidence has come from nonexperimental and longitudinal studies that investigate the assumptions of the reciprocal I/E model (cf., Marsh et al., 2017). In the reciprocal I/E model it is predicted that, within domains, achievement in math and achievement in the native language and the respective subject-specific academic self-concept have positive effects on subsequent achievement in math and achievement in the native language and subject-specific academic self-concept within domains but have negative effects across domains (e.g., Möller et al., 2014). The predictions of the model could be confirmed in several studies using path-analytic methods and nonexperimental data (see Möller et al., 2011; Möller et al., 2014; Niepel et al., 2014).

Despite the interplay between the three comparison processes, several aspects remain unclear within dimensional comparisons theory. This concern the benefits of dimensional comparison processes (see Müller-Kalthoff, Jansen et al., 2017). In addition, there is a need to investigate further moderators besides the grade difference between subjects of dimensional comparisons that are cognitive or motivational characteristics (Steinmayr & Spinath, 2015).

6.4.3 Stability and change: Sample level and individual level

A major limitation of Study 1 is that no predictors and outcome variables of change on group level and individual level were included in the analysis. Which variables predict changes in mean level? Which variables explain interindividual differences in intraindividual change? For instance, Schaffhuser et al. (2016) investigated the development of self-esteem and specific self-concept facets in the transition from late childhood to early adolescence (3 time points over 2 years; N = 248) and tested whether gender, puberty, and school transition explain interindividual differences in intraindividual change. Gender and school transition partly explained interindividual differences in intraindividual change in all self-concept facets. Girls were more negatively affected by school transition in comparison with boys.

Further, could there be groups of individuals showing different trajectories? This requires person-centered approaches like growth mixture analysis. Further, which variables (classroom context, personality, gender, age, socioeconomic status, person-environment fit) are predictive for different patterns of change and/or mean-level change? Moreover, are different groups differently associated with outcome variables like achievement, mental health?

Moreover, previous research showed that people with low self-esteem show an instable self-concept compared to those with high self-esteem (see Harter, 2012b). Is this also the case for academic self-concept or social self-concept? These remain questions that have to be answered in future research.

6.4.4 Causal interplay between self-concept facets

Given the limitation of the number of investigated self-concept facets outlined in the previous section, further self-concepts can be considered. The three self-concept areas were selected because they are significant for mastering developmental tasks in the social and academic area as well as in the development of self-esteem (see Chapter 1). However, adolescence is a time of major physical changes (i.e., puberty) leading to changes in appearance (i.e., body) and the interest in intimate relations with individuals. Therefore, physical self-concept may influence self-esteem. Thus, regarding the causal interplay between self-concept facets in adolescence, especially physical self-concept becomes salient with puberty. Harter (2012b) assumed that the effect of physical appearance on self-esteem is mediated by social acceptance in adolescence. There is some evidence supporting this assumption (Schmidt, Blum, Valkanover, & Conzelmann, 2015). Further, Morin et al. (2011) found bottom-up effects with respect to body image self-concept. In their study based on data from a 4-year, 6-wave, prospective longitudinal study of 1,001 adolescents, using autoregressive latent growth curve modeling, they found that statelike deviations in body image levels predicted later levels of self-esteem significantly but small in size whereas the relation between statelike deviations in self-esteem levels and later levels of body image was nonsignificant at all time points. As outlined in Section 6.2.1, future research should theoretically and empirically account for trait and state components within a particular self-concept facet. Further, to better compare research findings from different studies, analytic approaches that handle time in a continuous way seems appropriate (see Delsing & Oud, 2012; Oud & Voelkle, 2014; Voelkle et al., 2012). Further, the processes (i.e., self-strategies) behind the longitudinal interplay between self-concept facets should be given more attention in future studies (see Hattie, 2004).

7 Conclusions

Self-concepts are important variables that foster positive outcomes in different areas of life like social relations, academic career, and mental and physical health (e.g., Baumeister et al., 2003; Craven & Marsh, 2008; Trzesniewski et al., 2006). Fostering positive (and realistic) self-concepts is demanded as a major goal in education (Marsh & Hau, 2003b). The findings from the present dissertation contribute to the knowledge of different development aspects and structural assumptions and determinants of multiple self-concepts from childhood to late adolescence, which is a prerequisite to foster positive self-concepts in students (see also Harter, 2012b). Three vital self-concept areas (self-esteem, academic, and social self-concepts) were investigated.

It could be shown that a recently developed structural model of academic self-concept, the nested Marsh/Shavelson model (Brunner et al., 2010), demonstrated a good fit in elementary school students, which allows insights into children's profiles of strengths and weaknesses. Early in their school career, children show strengths and weaknesses in math and the native language. Further, dimensional (and social) comparisons are major determinants of subject-specific self-concepts as posited in the I/E model (Marsh, 1986) across different age groups and different statistical methods, and they are, for the most part, comparable over different structural models of academic self-concept. The size of the dimensional comparisons was moderated by the grade difference, that is, with growing grade difference in math and native languages, math and native language self-concepts become more distinct. Derived from these results, a practical implication for teachers is to keep in mind that dimensional comparisons influence their perceptions of students' ability/achievement in a specific school subject across different age ranges from middle childhood to late adolescence and, therefore, to provide achievement feedback as accurately as possible and therefore to promote that students get a profile of strength and weaknesses that are based on accurate criteria. Together with the finding from all studies that general academic self-concept declined in elementary school and secondary school on average in academic self-concept from late childhood to late adolescence, intervention programs and educational practices to foster positive general academic self-concepts should be carried out early, that is, right at the start of elementary school, to counteract a possible negative developmental trajectory in academic self-concept. Social self-concepts developed positive and self-esteem stayed constant throughout adolescence. Only some students showed a negative development.

Future research is needed to investigate dimensional comparisons in an experimental design and to investigate further moderators to support dimensional comparison theory to a theory based solidly on empirical truth. Further, the causal interplay between self-concept facets should be theoretically reconsidered by taking state and trait components of self-concepts into account. Finally, the underlying processes of developmental dynamics within a particular self-concept facet as well as the processes regarding the interplay of these are important for a comprehensive insight on self-perceptions which should be the focus of future self-concept research.

References

Aiken, L. R., & Groth-Marnat, G. (2006). *Psychological testing and assessment* (12. ed.). Boston: Allyn and Bacon.

Albert, S. (1977). Temporal comparison theory. *Psychological review*, *84*(6), 485–503. https://doi.org/10.1037/0033-295X.84.6.485

Aldrup, K., Klusmann, U., Lüdtke, O., Göllner, R., & Trautwein, U. (2018). Social support and classroom management are related to secondary students' general school adjustment: A multilevel structural equation model using student and teacher ratings. *Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1037/edu0000256

Anderman, E. M., & Midgley, C. (1997). Changes in achievement goal orientations, perceived academic competence, and grades across the transition to middle-level schools. *Contemporary Educational Psychology*, *22*(3), 269–298. https://doi.org/10.1006/ceps.1996.0926

Archambault, I., Eccles, J. S., & Vida, M. N. (2010). Ability self-concepts and subjective value in literacy: Joint trajectories from grades 1 through 12. *Journal of Educational Psychology*, *102*(4), 804–816. https://doi.org/10.1037/a0021075

Arens, A. K., Bodkin-Andrews, G., Craven, R. G., & Yeung, A. S. (2014). Self-concept of Indigenous and non-Indigenous Australian students: Competence and affect components and relations to achievement. *Learning and Individual Differences*, *32*, 93–103. https://doi.org/10.1016/j.lindif.2014.03.019

Arens, A. K., & Jansen, M. (2016). Self-concepts in reading, writing, listening, and speaking: A multidimensional and hierarchical structure and its generalizability across native and foreign languages. *Journal of Educational Psychology*, *108*(5), 646–664. https://doi.org/10.1037/edu0000081

Arens, A. K., Möller, J., & Watermann, R. (2016). Extending the internal/external frame of reference model to social studies: Self-concept and achievement in history and politics. *Learning and Individual Differences*, *51*, 91–99. https://doi.org/10.1016/j.lindif.2016.08.044

Arens, A. K., & Morin, A. J. S. (2016). Improved representation of the self-perception profile for children through bifactor exploratory structural equation modeling. *American Educational Research Journal*, *54*(1), 59–87. https://doi.org/10.3102/0002831216666490

Arens, A. K., Trautwein, U., & Hasselhorn, M. (2011). Erfassung des Selbstkonzepts im mittleren Kindesalter: Validierung einer deutschen Version des SDQ-I [Self-Concept Measurement with Preadolescent Children: Validation of a German Version of the SDQ I]. *Zeitschrift für Pädagogische Psychologie*, *25*(2), 131–144. https://doi.org/10.1024/1010-0652/a000030

Arens, A. K., Yeung, A. S., Craven, R. G., & Hasselhorn, M. (2013). A short German version of the Self Description Questionnaire I: Theoretical and empirical comparability. *International Journal of Research & Method in Education*, *36*(4), 415–438. https://doi.org/10.1080/1743727X.2012.710503

Arens, A. K., Yeung, A. S., & Hasselhorn, M. (2013). Native language self-concept and reading self-concept: Same or different? *The Journal of Experimental Education*, *82*(2), 229–252. https://doi.org/10.1080/00220973.2013.813362

Asparouhov, T., & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(3), 397–438. https://doi.org/10.1080/10705510903008204

Asparouhov, T., Muthén, B., & Morin, A. J. S. (2015). Bayesian structural equation modeling with cross-loadings and residual covariances. *Journal of Management*, *41*(6), 1561–1577. https://doi.org/10.1177/0149206315591075

Baudson, T. G., & Preckel, F. (2015). *Der FEESS-K. Eine Kurzversion des FEESS von Rauer und Schuck zum Einsatz in Klassenstufe 1 bis 4 [A short version of FEESS for grades 1 to 4]*. Unpublished instrument. University of Trier, Trier, Germany.

Baudson, T. G., Wollschläger, R., & Preckel, F. (2016). *Test zur Erfassung der Intelligenz im Kindesalter–THINK 1–4*. Göttingen: Hogrefe.

Baumeister, R. F. (Ed.). (1999). *The self in social psychology*. *Key readings in social psychology*. Philadelphia, PA: Psychology Press.

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: A journal of the American Psychological Society*, *4*(1), 1–44. https://doi.org/10.1111/1529-1006.01431

Becker, M., Lüdtke, O., Trautwein, U., Köller, O., & Baumert, J. (2012). The differential effects of school tracking on psychometric intelligence: Do academic-track schools make students smarter? *Journal of Educational Psychology*, *104*(3), 682–699. https://doi.org/10.1037/a0027608

Becker, M., & Neumann, M. (2017). Longitudinal big-fish-little-pond effects on academic self-concept development during the T transition from elementary to secondary schooling. *Journal of Educational Psychology*. Advance online publication. https://doi.org/10.1037/edu0000233

Berndt, T. J., & Burgy, L. (1996). Social self-concept. In B. A. Bracken (Ed.), *A Wiley-Interscience publication. Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 171–209). New York: Wiley.

Birkeland, M. S., Melkevik, O., Holsen, I., & Wold, B. (2012). Trajectories of global self-esteem development during adolescence. *Journal of adolescence*, *35*(1), 43–54. https://doi.org/10.1016/j.adolescence.2011.06.006

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*(3), 909–923. https://doi.org/10.2307/1131226

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. *Wiley series in probability and statistics*. Hoboken NJ u.a.: Wiley-Interscience.

Bolognini, Plancherel, Bettschart, & Halfon. (1996). Self-esteem and mental health in early adolescence: Development and gender differences. *Journal of adolescence*, *19*(3), 233–245.

Brown, J. D. (1993). Self-esteem and self-evaluation: Feeling is believing. In J. Suls (Ed.), *Psychological perspectives on the self, Volume 4: The self in social perspective* (pp. 27–58). Hillsdale. NJ: Erlbaum.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. *Methodology in the social sciences*. New York: Guilford Press.

Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., & Martin, R. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. *Journal of Educational Psychology*, *102*(4), 964–981. https://doi.org/10.1037/a0019644

Brunner, M., Keller, U., Hornung, C., Reichert, M., & Martin, R. (2009). The cross-cultural generalizability of a new structural model of academic self-concepts. *Learning and Individual Differences*, *19*(4), 387–403. https://doi.org/10.1016/j.lindif.2008.11.008

Brunner, M., Lüdtke, O., & Trautwein, U. (2008). The internal/external frame of reference model revisited: Incorporating general cognitive ability and general academic self-concept. *Multivariate behavioral research*, *43*(1), 137–172. https://doi.org/10.1080/00273170701836737

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of personality*, *80*(4), 796–846. https://doi.org/10.1111/j.1467-6494.2011.00749.x

Burnett, P. (2003). The impact of teacher feedback on student self-talk and self-concept in reading and mathematics. *The Journal of Classroom Interaction*, *38*(1), 11–16. Retrieved from http://www.jstor.org/stable/23874243.

Burnett, P. C. (2004). Enhancing students' self-perceptions: The impact of programs and teacher feedback. *Australian Journal of Guidance and Counselling*, *14*(1), 34–47. https://doi.org/10.1017/S103729110000265X

Byrne, B. M. (1984). The general/academic self-concept nomological network: A review of construct validation research. *Review of Educational Research, 54*(3), 427–456. https://doi.org/10.3102/00346543054003427

Byrne, B. M. (1986). Self-concept/academic achievement relations: An investigation of dimensionality, stability, and causality. *Canadian Journal of Behavioural Science / Revue canadienne des sciences du comportement*, *18*(2), 173–186. https://doi.org/10.1037/h0079982

Byrne, B. M. (1996). *Measuring self-concept across the life span: Issues and instrumentation.* Washington, DC: American Psychological Association.

Byrne, B. M., & Gavin, D. A. W. (1996). The Shavelson Model revisited: Testing for the structure of academic self-concept across pre-, early, and late adolescents. *Journal of Educational Psychology*, *88*(2), 215–228. https://doi.org/10.1037/0022-0663.88.2.215

Byrne, B. M., & Shavelson, R. J. (1996). On the structure of social self-concept for pre-, early-, and late adolescents: A test of the Shavelson, Hubner, and Stanton (1976) model. *Journal of Personality and Social Psychology, 70*(3), 599–613. https://doi.org/10.1037//0022-3514.70.3.599

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*(6), 561–570. https://doi.org/10.1080/01650250444000289

Chen, F. F., West, S. G., & Sousa, K. H. (2006). A comparison of bifactor and second-order models of quality of life. *Multivariate behavioral research*, *41*(2), 189–225. https://doi.org/10.1207/s15327906mbr4102_5

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233–255. https://doi.org/10.1207/s15328007sem0902_5

Christopher, J. S., Nangle, D. W., & Hansen, D. J. (1993). Social-skills interventions with adolescents: Current issues and procedures. *Behavior Modification*, *17*(3), 314–338. https://doi.org/10.1177/01454455930173005

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. https://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*(3), 469–483. https://doi.org/10.1037/a0035135

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.).* Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3. ed.). Mahwah, NJ: Erlbaum.

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*(6), 1723–1746. https://doi.org/10.1111/1467-8624.00375

Corcoran, K., Crusius, J., & Mussweiler, T. (2011). Social comparison: Motives, standards, and mechanisms. In D. Chadee (Ed.), *Theories in social psychology* (pp. 119–139). Oxford, UK: Wiley-Blackwell.

Craven, R. G., & Marsh, H. W. (2008). The centrality of the self-concept construct for psychological wellbeing and unlocking human potential: Implications for child and educational psychologists. *Educational & Child Psychology*, *25*(2), 104–118.

Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin*, *50*(6), 456–473. https://doi.org/10.1037/h0057173

Davies, J., & Brember, I. (1994). The Reliability and Validity of the 'Smiley' Scale. *British Educational Research Journal*, *20*(4), 447–454. https://doi.org/10.1080/0141192940200406

Delsing, M. J. M. H., & Oud, J. H. L. (2012). Causal directions between adolescents' externalising and internalising problems: A continuous-time analysis. *Netherlands Journal of Psychology*, *67*, 68–80.

Demo, D. (1992). The self-concept over time: Research issues and directions. *Annual Review of Sociology*, *18*(1), 303–326. https://doi.org/10.1146/annurev.soc.18.1.303

Dickhäuser, O., & Galfe, E. (2004). Besser als …, schlechter als… Leistungsbezogene Vergleichsprozesse in der Grundschule [Better …, worse … Achievement related comparison processes in elementary school]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *36*(1), 1–9. https://doi.org/10.1026/0049-8637.36.1.1

Dickhäuser, O., Reuter, M., & Hilling, C. (2005). Coursework selection: A frame of reference approach using structural equation modelling. *The British journal of educational psychology*, *75*(4), 673–688. https://doi.org/10.1348/000709905X37181

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*(2), 121–149. https://doi.org/10.1177/0748175610373459

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: Academic Press.

Dutton, K. A., & Brown, J. D. (1997). Global self-esteem and specific self-views as determinants of people's reactions to success and failure. *Journal of Personality and Social Psychology*, *73*(1), 139–148. https://doi.org/10.1037//0022-3514.73.1.139

Dweck, C. (2002). Development of achievement motivation. In A. Wigfield & J. Eccles (Eds.), *Development of Achievement Motivation* (pp. 57–88). Elsevier.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Iver, D. M. (1993). Development during adolescence. The impact of stage-environment fit on young adolescents' experiences in schools and in families. *The American psychologist*, *48*(2), 90–101.

Eccles, J. (2009). Who Am I and What Am I Going to Do With My Life? Personal and Collective Identities as Motivators of Action. *Educational Psychologist*, *44*(2), 78–89. https://doi.org/10.1080/00461520902832368

Eccles, J. S., & Roeser, R. W. (2009). Schools, Academic Motivation, and Stage-Environment Fit. In R. M. Lerner & L. D. Steinberg (Eds.), *Handbook of adolescent psychology,Volume 1: Individual Bases of Adolescent Development* (3rd ed.). Hoboken, NJ, Hoboken, NJ: Wiley Interscience.

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*(2), 283–310. https://doi.org/10.1111/j.1467-6494.1989.tb00484.x

Ehm, J.-H., Lindberg, S., & Hasselhorn, M. (2014). Reading, writing, and math self-concept in elementary school children: Influence of dimensional comparison processes. *European Journal of Psychology of Education*, *29*(2), 277–294. https://doi.org/10.1007/s10212-013-0198-x

Eid, M., Geiser, C., Koch, T., & Heene, M. (2017). Anomalous results in G-factor models: Explanations and alternatives. *Psychological methods*, *22*(3), 541–562. https://doi.org/10.1037/met0000083.

Eid, M., Lischetzke, T., Nussbeck, F. W., & Trierweiler, L. I. (2003). Separating trait effects from trait-specific method effects in multitrait-multimethod models: A multiple-indicator CT-C(M-1) model. *Psychological methods*, *8*(1), 38–60.

Elias, M. J. (1995). Primary prevention as health and social competence promotion. *The journal of primary prevention*, *16*(1), 5–24. https://doi.org/10.1007/BF02407230

Ellis, P. D. (2010). *The Essential Guide to Effect Sizes*. Cambridge: Cambridge University Press.

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford Publications.

Erikson, E. H. (1971). *Identity: Youth and crisis*. *Faber paper covered editions*. London: Faber & Faber.

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*(3), 607–619. https://doi.org/10.1037/a0024299

Esnaola, I., Elosua, P., & Freeman, J. (2018). Internal structure of academic self-concept through the Self-Description Questionnaire II-Short (SDQII-S). *Learning and Individual Differences*, *62*, 174–179. https://doi.org/10.1016/j.lindif.2018.02.006

European Commission/EACEA/Eurydice. (2017). *The Structure of the European Education Systems 2017/18: Schematic Diagrams: Eurydice Facts and Figures*. Luxembourg: Publications Office of the European Union. DOI: 10.2797/2705.

Felson, R. B. (1989). Parents and the reflected appraisal process: A longitudinal analysis. *Journal of Personality and Social Psychology*, *56*(6), 965–971. https://doi.org/10.1037//0022-3514.56.6.965

Fend, H., Georg, W., Berger, A., Grob, U., & Lauterbach, W. (2002). *Lebensverläufe von der späten Kindheit ins frühe Erwachsenenalter (LifE). Die Bedeutung von Erziehungserfahrungen und Entwicklungsprozessen für die Lebensbewältigung [Pathways from late childhood to adulthood. Context and development in adolescence as predictors of productive life-courses]. Unpublished manuscript. University of Zurich, and University of Constance, Zurich, Switzerland, Constance, Germany.* Unpublished manuscript. University of Zurich, and University of Constance, Zurich, Switzerland, Constance, Germany.

Fend, H.-G., & Prester, H. (1986). *Dokumentation der Skalen des Projekts "Entwicklung im Jugendalter". Bericht aus dem Projekt "Entwicklung im Jugendalter" (T. 3).* Konstanz, Universität Konstanz, Sozialwissenschaftliche Fakultät.

Fergusson, D. M., & Woodward, L. J. (2002). Mental health, educational, and social role outcomes of adolescents with depression. *Archives of General Psychiatry*, *59*(3), 225. https://doi.org/10.1001/archpsyc.59.3.225

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Filipp, S.-H., & Mayer, A.-K. (2005). Selbstkonzept-Entwicklung. In *J. B. Asendorpf [Hrsg.] Enzyklopädie der Psychologie: Entwicklungspsychologie: Soziale, emotionale und Persönlichkeitsentwicklung. Göttingen: Hogrefe* (pp. 259–334).

Fraine, B. de, van Damme, J., & Onghena, P. (2007). A longitudinal analysis of gender differences in academic self-concept and language achievement: A multivariate multilevel latent growth approach. *Contemporary Educational Psychology*, *32*(1), 132–150. https://doi.org/10.1016/j.cedpsych.2006.10.005

Fredricks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental psychology*, *38*(4), 519–533. https://doi.org/10.1037/0012-1649.38.4.519

Fryer, J. W., & Elliot, A. J. (2007). Stability and change in achievement goals. *Journal of Educational Psychology*, *99*(4), 700–714. https://doi.org/10.1037/0022-0663.99.4.700

Geiser, C., Keller, B., & Lockhart, G. (2013). First versus second order latent growth curve models: Some insights from latent state-trait theory. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(3). https://doi.org/10.1080/10705511.2013.797832

Gibbons, F. X., Lane, D. J., Gerrard, M., Reis-Bergan, M., Lautrup, C. L., Pexa, N. A., & Blanton, H. (2002). Comparison-level preferences after performance: Is downward comparison theory still useful? *Journal of Personality and Social Psychology*, *83*(4), 865–880. https://doi.org/10.1037/0022-3514.83.4.865

Gogol, K., Brunner, M., Martin, R., Preckel, F., & Goetz, T. (2017). Affect and motivation within and between school subjects: Development and validation of an integrative structural model of academic self-concept, interest, and anxiety. *Contemporary Educational Psychology*, *49*, 46–65. https://doi.org/10.1016/j.cedpsych.2016.11.003

Gogol, K., Brunner, M., Preckel, F., Goetz, T., & Martin, R. (2016). Developmental dynamics of general and school-subject-specific components of academic self-concept, academic interest, and academic anxiety. *Frontiers in psychology*, *7*, 356. https://doi.org/10.3389/fpsyg.2016.00356

Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*(1), 26–37.

Granleese, J., & Joseph, S. (1994). Reliability of the Harter Self-Perception Profile for children and predictors of global self-worth. *The Journal of genetic psychology*, *155*(4), 487–492. https://doi.org/10.1080/00221325.1994.9914796

Guay, F., Larose, S., & Boivin, M. (2004). Academic self-concept and educational attainment Level: A ten-year longitudinal study. *Self and Identity*, *3*(1), 53–68. https://doi.org/10.1080/13576500342000040

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, *95*(1), 124–136. https://doi.org/10.1037/0022-0663.95.1.124

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research*, *52*(1), 123–142. https://doi.org/10.3102/00346543052001123

Harris, M. A., Donnellan, M. B., & Trzesniewski, K. H. (2017). The lifespan self-esteem scale: Initial validation of a new measure of global self-esteem. *Journal of personality assessment*, 1–12. https://doi.org/10.1080/00223891.2016.1278380

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: Guilford Press.

Harter, S. (2005). Self-concepts and self-esteem, children and adolescents. In C. B. Fisher & R. M. Lerner (Eds.), Encyclopedia of Applied Developmental Science (pp. 973–977).California; United States: SAGE Publications, Inc. https://doi.org/10.4135/9781412950565.n368

Harter, S. (2012a). *Self-perception profile for adolescents: Manual and questionaires (Revision 2012)*. Retrieved from https://portfolio.du.edu/SusanHarter/page/44210.

Harter, S. (2012b). *The construction of the self: Developmental and sociocultural foundations* (2. ed.). New York, NY: Guilford Press.

Hattie, J. (1992). *Self-concept*. Hillsdale [New Jersey]: Lawrence Erlbaum.

Hattie, J. (2004). Models of self-concept that are neither top-down or bottom-up: The rope model of self-concept [R]. Retrieved from https://cdn.auckland.ac.nz/assets/education/hattie/docs/hattie-models-of-self-concept-(2004).pdf

Havighurst, R. J. (1979). *Developmental tasks and education* (3rd ed., 5th print). New York: Longman.

Hawley, P. H. (2003). Prosocial and coercive configurations of resource control in early adolescence: A Case for the Well-Adapted Machiavellian. *Merrill-Palmer Quarterly*, *49*(3), 279–309. https://doi.org/10.1353/mpq.2003.0013

Helm, F., Mueller-Kalthoff, H., Nagy, N., & Möller, J. (2016). Dimensional Comparison Theory: Perceived subject similarity impacts on Students' self-concepts. *AERA Open*, *2*(2), 1-9. https://doi.org/10.1177/2332858416650624

Howard, M. C., & Hoffman, M. E. (2017). Variable-centered, person-centered, and person-specific approaches. *Organizational Research Methods, 19*(1), 109442811774402. https://doi.org/10.1177/1094428117744021

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://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*(3), 251–260. https://doi.org/10.1037/a0020543

Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of school psychology*, *49*(5), 505–528. https://doi.org/10.1016/j.jsp.2011.07.001

Huitt, W. (2011). Self and self-views. *Educational Psychology Interactive.* Retrieved from http://www.edpsycinteractive.org/topics/self/self.html

IBM Corp. (2014). IBM SPSS statistics for windows, version 23.0. Armonk, NY: IBM Corp.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of consulting and clinical psychology*, *59*(1), 12–19.

James, W. (1890). *The principles of psychology* (Bd. 1). New York: Holt.

James, W. (1892). *Psychology: The briefer course*. New York: Holt.

Jansen, M., Lüdtke, O., & Schroeders, U. (2016). Evidence for a positive relation between interest and achievement: Examining between-person and within-person variation in five domains. *Contemporary Educational Psychology*, *46*, 116–127. https://doi.org/10.1016/j.cedpsych.2016.05.004

Jansen, M., Schroeders, U., Lüdtke, O., & Marsh, H. W. (2015). Contrast and assimilation effects of dimensional comparisons in five subjects: An extension of the I/E model. *Journal of Educational Psychology*, *107*(4), 1086–1101. https://doi.org/10.1037/edu0000021

Jonkmann, K., Trautwein, U., & Lüdtke, O. (2009). Social dominance in adolescence: The moderating role of the classroom context and behavioral heterogeneity. *Child Development*, *80*(2), 338–355. https://doi.org/10.1111/j.1467-8624.2009.01264.x

Ju, S., & Lee, Y. (2017). Developmental trajectories and longitudinal mediation effects of self-esteem, peer attachment, child maltreatment and depression on early adolescents. *Child abuse & neglect*, *76*, 353–363. https://doi.org/10.1016/j.chiabu.2017.11.015

Kernis, M. H. (1993). The roles of stability and level of self-esteem in psychological Functioning. In R. F. Baumeister (Ed.), *Self-Esteem* (pp. 167–182). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4684-8956-9_9

Kernis, M. H. (Ed.). (2006). *Self-Esteem: Issues and answers*. A Sourcebook of Current Perspectives. New York, NY: Psychology Press.

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (Fourth edition). *Methodology in the social sciences*. New York: The Guilford Press.

Krajewski, K., Nieding, G., & Schneider, W. (2008). Kurz- und langfristige Effekte mathematischer Frühförderung im Kindergarten durch das Programm „Mengen, zählen, Zahlen”. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *40*(3), 135–146. https://doi.org/10.1026/0049-8637.40.3.135

Krapp, K. (2005). *The Gale encyclopedia of children's health: Infancy through adolescence*. Detroit: Thomson Gale.

Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M.,. . . Weiß, M. (2002). *PISA 2000: Dokumentation der Erhebungsinstrumente*. *Materialien aus der Bildungsforschung: Nr. 72*. Berlin: Max-Planck-Inst. für Bildungsforschung.

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. *The Journal of Early Adolescence*, *34*(5), 638–666. https://doi.org/10.1177/0272431613507498

Lau, I. C.-y., Yeung, A. S., Jin, P., & Low, R. (1999). Toward a hierarchical, multidimensional English self-concept. *Journal of Educational Psychology*, *91*(4), 747–755. https://doi.org/10.1037//0022-0663.91.4.747

Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development*, *31*(4), 357–365. https://doi.org/10.1177/0165025407077757

Little, T. D., Slegers, D. W., & Card, N. A. (2006). A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling: A Multidisciplinary Journal*, *13*(1), 59–72. https://doi.org/10.1207/s15328007sem1301_3

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A Social Psychological Perspective. *Annual review of psychology*, *38*(1), 299–337. https://doi.org/10.1146/annurev.ps.38.020187.001503

Marsh, H. W. (1986). Verbal and Math Self-Concepts: An internal/external frame of reference Model. *American Educational Research Journal*, *23*(1), 129–149. https://doi.org/10.3102/00028312023001129

Marsh, H. W. (1988). *Self Description Questionnaire: A theoretical and empirical basis for the measurement of multiple dimensions of preadolescent self-concept: A test manual and a research monograph.* San Antonio, TX: Psychological Corporation.

Marsh, H. W. (1990a). *Self Description Questionaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept: An interim test manual and a research monograph.* San Antonio, TX: The Psychological Corporation.

Marsh, H. W., Craven, R. G., & Martin, A. J. (2006). What is the nature of self-esteem? Unidimensional and multidimensional perspectives. In M. H. Kernis (Ed.), *Self-Esteem: Issues and answers. A Sourcebook of Current Perspectives.* New York, NY: Psychology Press.

Marsh, H. W., & Hattie, J. (1996). Theoretical perspectives of the structure of self-concept. In B. A. Bracken (Ed.), *A Wiley-Interscience publication. Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 38–90). New York: Wiley.

Marsh, H. W., & Yeung, A. S. (1997). Coursework Selection: Relations to Academic Self-Concept and Achievement. *American Educational Research Journal*, *34*(4), 691–720. https://doi.org/10.3102/00028312034004691

Marsh, H. W., Xu, M., & Martin, A. J. (2012). Self-concept: A synergy of theory, method, and application. In K. R. Harris (Ed.), *APA educational psychology handbook.Volume 1: Theories, Constructs, and Critical Issues* (pp. 427–458). Washington DC: American Psychological Association.

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, *79*(3), 280–295. https://doi.org/10.1037/0022-0663.79.3.280

Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology*, *81*(3), 417–430. https://doi.org/10.1037/0022-0663.81.3.417

Marsh, H. W. (1990b). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, *2*(2), 77–172. https://doi.org/10.1007/bf01322177

Marsh, H. W. (1990c). Influences of internal and external frames of reference on the formation of math and English self-concepts. *Journal of Educational Psychology*, *82*(1), 107–116. https://doi.org/10.1037/0022-0663.82.1.107

Marsh, H. W. (1990d). *Self Description Questionnaire-III: SDQ III Manual*: University of Western Sydney, Macarthur.

Marsh, H. W. (1990e). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology*, *82*(4), 623–636. https://doi.org/10.1037/0022-0663.82.4.623

Marsh, H. W. (1993a). Academic self-concept: Theory, meaurement, and research. In J. Suls (Ed.), *Psychological perspectives on the self, Volume 4: The self in social perspective* (pp. 59–98). Hillsdale. NJ: Erlbaum.

Marsh, H. W. (1993b). Relations between global and specific domains of self: The importance of individual importance, certainty, and ideals. *Journal of Personality and Social Psychology*, *65*(5), 975–992. https://doi.org/10.1037//0022-3514.65.5.975

Marsh, H. W. (2005). Big-fish-little-pond effect on academic self-concept. *Zeitschrift für Pädagogische Psychologie*, *19*(3), 119–129. https://doi.org/10.1024/1010-0652.19.3.119

Marsh, H. W., & Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *Journal of Educational Psychology*, *95*(4), 687–706. https://doi.org/10.1037/0022-0663.95.4.687

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*(5), 940–956. https://doi.org/10.1037/0022-0663.76.5.940

Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, *80*(3), 366–380. https://doi.org/10.1037//0022-0663.80.3.366

Marsh, H. W., & Craven, R. (1997). Academic Self-Concept: Beyond The Dustbowl. In G. D. Phye (Ed.), *Educational psychology series. Handbook of classroom assessment: Learning, achievement, and adjustment* (pp. 131–198). San Diego: Academic Press.

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, *1*(2), 133–163. https://doi.org/10.1111/j.1745-6916.2006.00010.x

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*(3), 377–392. https://doi.org/10.1037/0022-0663.83.3.377

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*(4), 1030. https://doi.org/10.2307/1132361

Marsh, H. W., Ellis, L. A., & Craven, R. G. (2002). How do preschool children feel about themselves? Unraveling measurement and multidimensional self-concept structure. *Developmental psychology*, *38*(3), 376–393. https://doi.org/10.1037//0012-1649.38.3.376

Marsh, H. W., & Hau, K.-T. (2003a). Big-fish-little-pond effect on academic self-concept. A cross-cultural (26-country) test of the negative effects of academically selective schools. *The American psychologist*, *58*(5), 364–376.

Marsh, H. W., & Hau, K.-T. (2003b). Big-fish-little-pond effect on academic self-concept. A cross-cultural (26-country) test of the negative effects of academically selective schools. *The American psychologist*, *58*(5), 364–376.

Marsh, H. W., & Hau, K.-T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal/external frame of reference predictions Across 26 Countries. *Journal of Educational Psychology*, *96*(1), 56–67. https://doi.org/10.1037/0022-0663.96.1.56

Marsh, H. W., Hau, K.-T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's Brief Self-Report Measure of Educational Psychology's Most Useful Affective Constructs: Cross-Cultural, Psychometric Comparisons Across 25 Countries. *International Journal of Testing*, *6*(4), 311–360. https://doi.org/10.1207/s15327574ijt0604_1

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling: A Multidisciplinary Journal*, *11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, *97*(3), 562–582. https://doi.org/10.1037/0033-2909.97.3.562

Marsh, H. W., Kong, C.-K., & Hau, K.-T. (2001). Extension of the internal/external frame of reference model of self-concept formation: Importance of native and nonnative languages for Chinese students. *Journal of Educational Psychology*, *93*(3), 543–553. https://doi.org/10.1037/0022-0663.93.3.543

Marsh, H. W., Kuyper, H., Seaton, M., Parker, P. D., Morin, A. J.S., Möller, J., & Abduljabbar, A. S. (2014). Dimensional comparison theory: An extension of the internal/external frame of reference effect on academic self-concept formation. *Contemporary Educational Psychology*, *39*(4), 326–341. https://doi.org/10.1016/j.cedpsych.2014.08.003

Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Abduljabbar, A. S., Abdelfattah, F., & Jansen, M. (2015). Dimensional Comparison Theory: Paradoxical relations between self-beliefs and achievements in multiple domains. *Learning and Instruction*, *35*, 16–32. https://doi.org/10.1016/j.learninstruc.2014.08.005

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *The British journal of educational psychology*, *81*(Pt 1), 59–77. https://doi.org/10.1348/000709910x503501

Marsh, H. W., Martin, A. J., Yeung, A. S., & Craven, R. G. (2017). Competence self-perceptions. In A. J. Elliot, C. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation: Theory and application* (pp. 85–115). New York, London, New York, London: The Guilford Press.

Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality & social psychology bulletin*, *34*(4), 542–552. https://doi.org/10.1177/0146167207312313

Marsh, H. W., & Seaton, M. (2015). The big-fish–little-pond effect, competence self-perceptions, and relativity. In *Advances in Motivation Science* (Vol. 2, pp. 127–184). Elsevier. https://doi.org/10.1016/bs.adms.2015.05.002

Marsh, H. W., & Shavelson, R. (1985). Self-Concept: Its Multifaceted, Hierarchical Structure. *Educational Psychologist*, *20*(3), 107–123. https://doi.org/10.1207/s15326985ep2003_1

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, *76*(2), 397–416. https://doi.org/10.1111/j.1467-8624.2005.00853.x

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Integration of multidimensional self-concept and core personality constructs: Construct validation and relations to well-being and achievement. *Journal of personality*, *74*(2), 403–456. https://doi.org/10.1111/j.1467-6494.2005.00380.x

Marsh, H. W., & Yeung, A. S. (1998). Top-down, bottom-up, and horizontal models: The direction of causality in multidimensional, hierarchical self-concept models. *Journal of Personality and Social Psychology*, *75*(2), 509–527. https://doi.org/10.1037//0022-3514.75.2.509

Marsh, H. W., Hau, K.-T., & Grayson, D. (2005). Goodness of fit in structural equation models. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Multivariate applications series. Contemporary psychometrics: A festschrift for Roderick P. McDonald. Contemporary psychometrics: A*

Festschrift for Roderick P. McDonald (pp. 275–340). Mahwah, NJ: Lawrence Erlbaum Associates.

Marx, R. W., & Winne, P. H. (1980). Self-validation research: Some current complexities. *Measurement and Evaluation in Guidance, 13*(2).

McArdle, J. J., & Nesselroade, J. R. (Eds.). (2014). *Longitudinal data analysis using structural equation models*. Washington, D.C.: American Psychological Association.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah NJ u.a.: Erlbaum.

McElhaney, K. B., Antonishak, J., & Allen, J. P. (2008). "They like me, they like me not": Popularity and adolescents' perceptions of acceptance predicting social functioning over time. *Child Development, 79*(3), 720–731. https://doi.org/10.1111/j.1467-8624.2008.01153.x.

Mellenbergh, G. J. (1989). Item bias and item response theory. *International Journal of Educational Research, 13*(2), 127–143. https://doi.org/10.1016/0883-0355(89)90002-5

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*(4), 525–543. https://doi.org/10.1007/BF02294825

Miller, S. A. (1998). *Developmental research methods (2nd ed.)*. Upper Saddle River, NJ: Prentice-Hall.

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate behavioral research, 39*(3), 479–515. https://doi.org/10.1207/S15327906MBR3903_4

Möller, J. (2016). The generalized internal/external frame of reference model: An extension to dimensional comparison theory. *Frontline Learning Research, 4*(2), 1–11. https://doi.org/10.14786/flr.v4i2.169

Möller, J., & Husemann, N. (2006). Internal comparisons in everyday life. *Journal of Educational Psychology, 98*(2), 342–353. https://doi.org/10.1037/0022-0663.98.2.342

Möller, J., & Köller, O. (2001). Dimensional comparisons: An experimental approach to the internal/external frame of reference model. *Journal of Educational Psychology, 93*(4), 826–835. https://doi.org/10.1037/0022-0663.93.4.826

Möller, J., & Köller, O. (2004). Die Genese akademischer Selbstkonzepte [On the development of academic self-concepts: The impact of social and dimensional comparisons]. *Psychologische Rundschau*, *55*(1), 19–27. https://doi.org/10.1026/0033-3042.55.1.19

Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological review*, *120*(3), 544–560. https://doi.org/10.1037/a0032459

Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, *79*(3), 1129–1167. https://doi.org/10.3102/0034654309337522

Möller, J., Pohlmann, B., Streblow, L., & Kaufmann, J. (2002). Die Spezifität von Begabungsüberzeugungen als Determinante des verbalen und mathematischen Begabungsselbstkonzepts. [Specificity of Ability Beliefs as Determinant of Academic Self-Concepts.]. *Zeitschrift für Pädagogische Psychologie*, *16*(2), 87–97. https://doi.org/10.1024//1010-0652.16.2.87

Möller, J., Retelsdorf, J., Köller, O., & Marsh, H. W. (2011). The reciprocal internal/external frame of reference model: An integration of models of relations between academic achievement and self-concept. *American Educational Research Journal*, *48*(6), 1315–1346. https://doi.org/10.3102/0002831211419649

Möller, J., & Savyon, K. (2003). Not very smart, thus moral: Dimensional comparisons between academic self-concept and honesty. *Social Psychology of Education*, *6*(2), 95–106. https://doi.org/10.1023/A:1023247910033

Möller, J., Streblow, L., & Pohlmann, B. (2006). The belief in a negative interdependence of math and verbal abilities as determinant of academic self-concepts. *The British journal of educational psychology*, *76*(Pt 1), 57–70. https://doi.org/10.1348/000709905X37451

Möller, J., Streblow, L., Pohlmann, B., & Köller, O. (2006). An extension to the internal/external frame of reference model to two verbal and numerical domains. *European Journal of Psychology of Education*, *21*(4), 467–487. https://doi.org/10.1007/BF03173515

Möller, J., Zimmermann, F., & Köller, O. (2014). The reciprocal internal/external frame of reference model using grades and test scores. *The British journal of educational psychology*, *84*(Pt 4), 591–611. https://doi.org/10.1111/bjep.12047

Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2015). A bifactor exploratory structural equation modeling framework for the identification of Distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(1), 116–139. https://doi.org/10.1080/10705511.2014.961800

Morin, A. J. S., Boudrias, J.-S., Marsh, H. W., McInerney, D. M., Dagenais-Desmarais, V., Madore, I., & Litalien, D. (2017). Complementary variable- and person-centered approaches to the dimensionality of psychometric constructs: Application to Psychological Wellbeing at Work. *Journal of Business and Psychology*, *32*(4), 395–419. https://doi.org/10.1007/s10869-016-9448-7

Morin, A. J. S., Maïano, 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*(2), 157–201. https://doi.org/10.1080/00273171.2010.546731

Morin, A. J. S., Maïano, C., Marsh, H. W., Nagengast, B., & Janosz, M. (2013). School life and adolescents' self-esteem trajectories. *Child Development*, *84*(6), 1967–1988. https://doi.org/10.1111/cdev.12089

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

Müller-Kalthoff, H., Helm, F., & Möller, J. (2017). The big three of comparative judgment: On the effects of social, temporal, and dimensional comparisons on academic self-concept. *Social Psychology of Education*, *20*(4), 849–873. https://doi.org/10.1007/s11218-017-9395-9

Müller-Kalthoff, H., Jansen, M., Schiefer, I. M., Helm, F., Nagy, N., & Möller, J. (2017). A double-edged sword? On the benefit, detriment, and net effect of dimensional comparison on self-concept. *Journal of Educational Psychology*, *109*(7), 1029–1047. https://doi.org/10.1037/edu0000171

Muthén, B. O., Du Toit, S. H.C., & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Retrieved from http://gseis.ucla.edu/faculty/muthen/articles/Article_075.pdf

Muthén, B. O., & Muthén, L. K. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

Muthén, B., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological methods*, *17*(3), 313–335. https://doi.org/10.1037/a0026802.

Nagengast, B., & Marsh, H. W. (2012). Big fish in little ponds aspire more: Mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *Journal of Educational Psychology*, *104*(4), 1033–1053. https://doi.org/10.1037/a0027697

Nangle, D. W., Hansen, D. J., Erdley, C. A., & Norton, P. J. (2009). *Practitioner's Guide to Empirically Based Measures of Social Skills*: Springer New York. Retrieved from https://books.google.de/books?id=g_b0CLDtlAUC

Nesselroade, J. R. (1992). Interindividual differences in intraindividual change. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (2nd ed., pp. 92–105). Washington, DC: American Psychological Assoc. https://doi.org/10.1037/10099-006

Newsom, J. T. (2015). *Longitudinal structural equation modeling: A comprehensive introduction*. *Multivariate applications series*. New York, NY: Routledge.

Nicholls, J. G. (1979). Development of perception of own attainment and causal attributions for success and failure in reading. *Journal of Educational Psychology, 71*(1), 94–99. https://doi.org/10.1037/0022-0663.71.1.94

Niepel, C., Brunner, M., & Preckel, F. (2014). The longitudinal interplay of students’ academic self-concepts and achievements within and across domains: Replicating and extending the reciprocal internal/external frame of reference model. *Journal of Educational Psychology*, *106*(4), 1170–1191. https://doi.org/10.1037/a0036307

Nussbeck, F. W., Eid, M., Geiser, C., Courvoisier, D. S., & Lischetzke, T. (2009). A CTC(M −1) Model for Different Types of Raters. *Methodology*, *5*(3), 88–98. https://doi.org/10.1027/1614-2241.5.3.88

OECD. (2003). *Student engagement at school: A sense of belonging and participation: Results from PISA 2000*. Paris: OECD. Retrieved from http://www.oecd.org/education/school/programmeforinternationalstudentassessmentpisa/33689437.pdf.

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do Self-Concept Interventions Make a Difference? A Synergistic Blend of Construct Validation and Meta-Analysis. *Educational Psychologist*, *41*(3), 181–206. https://doi.org/10.1207/s15326985ep4103_4

Orth, U., & Robins, R. W. (2014). The development of self-esteem. *Current Directions in Psychological Science*, *23*(5), 381–387. https://doi.org/10.1177/0963721414547414

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*(3), 695–708. https://doi.org/10.1037/0022-3514.95.3.695

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*(6), 1271–1288. https://doi.org/10.1037/a0025558

Oud, J. H. L., & Voelkle, M. C. (2014). Do missing values exist? Incomplete data handling in cross-national longitudinal studies by means of continuous time modeling. *Quality & Quantity*, *48*(6), 3271–3288. https://doi.org/10.1007/s11135-013-9955-9

Parker, P. D., Marsh, H. W., Morin, A. J. S., Seaton, M., & van Zanden, B. (2015). If one goes up the other must come down: Examining ipsative relationships between math and English self-concept trajectories across high school. *British Journal of Educational Psychology*, *85*(2), 172–191. https://doi.org/10.1111/bjep.12050

Pesu, L. A., Aunola, K., Viljaranta, J., & Nurmi, J.-E. (2016). The development of adolescents' self-concept of ability through grades 7-9and the role of parental beliefs. *Frontline Learning Research*, *4*(3), 92–109. https://doi.org/10.14786/flr.v4i3.249

Piaget, J., Inhelder, B., Sinclair-De Zwart, H., & Pomerans, A. J. (1979). *Memory and intelligence*. London: Routledge & Kegan Paul.

Pinxten, M., Wouters, S., Preckel, F., Niepel, C., Fraine, B. de, & Verschueren, K. (2015). The formation of academic self-concept in elementary education: A unifying model for external and internal comparisons. *Contemporary Educational Psychology*, *41*, 124–132. https://doi.org/10.1016/j.cedpsych.2014.12.003

Poloczek, S., Karst, K., Praetorius, A.-K., & Lipowsky, F. (2011). Generalisten oder Spezialisten? -Bereichsspezifität und leistungs-bezogene Zusammenhänge des schulischen Selbstkonzepts von Schulanfängern [Generalists or Specialists? Domain Specificity and Relationships to Achievement Data of the Academic Self-Concept of School Starters]. *Zeitschrift für Pädagogische Psychologie*, *25*(3), 173–183. https://doi.org/10.1024/1010-0652/a000045

Praetorius, A.-K., Kastens, C., Hartig, J., & Lipowsky, F. (2016). Haben Schüler mit optimistischen Selbsteinschätzungen die Nase vorn? [Are Students With Optimistic Self-Concepts One Step Ahead? Relations Between Optimistic, Realistic, and Pessimistic Self-Concepts and the Achievement Development of Primary School Children]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *48*(1), 14–26. https://doi.org/10.1026/0049-8637/a000140

Preckel, F., Niepel, C., Schneider, M., & Brunner, M. (2013). Self-concept in adolescence: A longitudinal study on reciprocal effects of self-perceptions in academic and social domains. *Journal of adolescence*, *36*(6), 1165–1175. https://doi.org/10.1016/j.adolescence.2013.09.001

Preckel, F., Zeidner, M., Goetz, T., & Schleyer, E. J. (2008). Female 'big fish' swimming against the tide: The 'big-fish-little-pond effect' and gender-ratio in special gifted classes. *Contemporary Educational Psychology*, *33*(1), 78–96. https://doi.org/10.1016/j.cedpsych.2006.08.001

Prinstein, M. J., & La Greca, A. M. (2002). Peer Crowd Affiliation and Internalizing Distress in Childhood and Adolescence: A Longitudinal Follow-Back Study. *Journal of Research on Adolescence*, *12*(3), 325–351. https://doi.org/10.1111/1532-7795.00036

Rauer, W., & Schuck, K.-D. (2003). *Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern dritter und vierter Klassen (FEESS 3-4)*. Göttingen: Beltz Tests.

Rauer, W., & Schuck, K.-D. (2004). *Fragebogen zur Erfassung emotionaler und sozialer Schulerfahrungen von Grundschulkindern ersten und zweiten Klassen (FEESS 1-2)*. Göttingen: Beltz Tests.

Riding, R. J., & Rayner, S. (Eds.). (2001). *Self perception*. *International perspectives on individual differences: v. 2*. Westport, Conn.: Ablex Pub.

Rieger, S., Göllner, R., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). Social Cognitive Constructs Are Just as Stable as the Big Five Between Grades 5 and 8. *AERA Open*, *3*(3), 233285841771769. https://doi.org/10.1177/2332858417717691

Rieger, S., Göllner, R., Trautwein, U., & Roberts, B. W. (2016). Low self-esteem prospectively predicts depression in the transition to young adulthood: A replication of Orth, Robins, and Roberts (2008). *Journal of Personality and Social Psychology*, *110*(1), e16-22. https://doi.org/10.1037/pspp0000037

Roberts, B. W., Wood, D., & Caspi, A. (2010). The development of personality traits in adulthood. In O. P. John, L. A. Pervin, & R. W. Robins (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 375–399). New York [etc]: Guilford Press.

Robins, R. W., & Trzesniewski, K. H. (2005). Self-esteem development across the lifespan. *Current Directions in Psychological Science*, *14*(3), 158–162. https://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*(3), 423–434. https://doi.org/10.1037//0882-7974.17.3.423

Robinson, K., Schmidt, T., & Teti, D. M. (2005). Issues in the use of longitudinal and cross-sectional designs. In D. M. Teti (Ed.), *Handbook of Research Methods in Developmental Science* (pp. 1–20). Oxford, UK: Blackwell Publishing Ltd. https://doi.org/10.1002/9780470756676.ch1

Rose-Krasnor, L. (1997). The nature of social Competence: A theoretical review. *Social Development*, *6*(1), 111–135. https://doi.org/10.1111/j.1467-9507.1997.tb00097.x

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

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review*, *60*(1), 141. https://doi.org/10.2307/2096350

Rost, D. H., Dickhäuser, O., Sparfeldt, J. R., & Schilling, S. R. (2004). Fachspezifische Selbstkonzepte und Schulleistungen im dimensionalen Vergleich. [Academic Self-Concepts, Achievement and the Internal Frame of Reference: A Design-Oriented Demonstration of Predictions Derived from the I/E-model]. *Zeitschrift für Pädagogische Psychologie*, *18*(1), 43–52. https://doi.org/10.1024/1010-0652.18.1.43

Rost, D. H., Sparfeldt, J. R., Dickhäuser, O., & Schilling, S. R. (2005). Dimensional comparisons in subject-specific academic self-concepts and achievements: A quasi-experimental approach. *Learning and Instruction*, *15*(6), 557–570. https://doi.org/10.1016/j.learninstruc.2005.08.003

Salvia, J., Sheare, J. B., & Algozzine, B. (1976). Facial attractiveness and personal-social development. *Journal of Abnormal Child Psychology*, *3*(3), 171–178. https://doi.org/10.1007/BF00916748

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment*, *29*(4), 347–363. https://doi.org/10.1177/0734282911406661

Sass, D. A., & Schmitt, T. A. (2013). Testing measurement and structural Invariance. In T. Teo (Ed.), *Handbook of quantitative methods for educational research* (pp. 315–345). Rotterdam: Sense. https://doi.org/10.1007/978-94-6209-404-8_15

Schaffhuser, K., Allemand, M., & Schwarz, B. (2016). The Development of Self-Representations During the Transition to Early Adolescence: The Role of Gender, Puberty, and School Transition. *The Journal of Early Adolescence*, *37*(6), 774–804. https://doi.org/10.1177/0272431615624841

Schilling, S. R., Sparfeldt, J. R., & Rost, D. H. (2004). Wie generell ist das Modell? Analysen zum Geltungsbereich des “Internal/External Frame of Reference”-Modells [How General is the Model? Analyzing the Applicability of the Internal/External Frame of Reference Model to More Than Two School Subjects.]. *Zeitschrift für Pädagogische Psychologie*, *18*(3/4), 221–230. https://doi.org/10.1024/1010-0652.18.34.221

Schmidt, M., Blum, M., Valkanover, S., & Conzelmann, A. (2015). Motor ability and self-esteem: The mediating role of physical self-concept and perceived social acceptance. *Psychology of Sport and Exercise*, *17*, 15–23. https://doi.org/10.1016/j.psychsport.2014.11.006

Schurtz, I. M., Pfost, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction*, *34*, 32–41. https://doi.org/10.1016/j.learninstruc.2014.08.001

Seeshing Yeung, A., Chui, H.-S., Lau, I. C.-y., McInerney, D. M., Russell-Bowie, D., & Suliman, R. (2000). Where is the hierarchy of academic self-concept? *Journal of Educational Psychology*, *92*(3), 556–567. https://doi.org/10.1037//0022-0663.92.3.556

Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. P. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 265–278). New York, NY: Guilford Press.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, *46*(3), 407–441. https://doi.org/10.3102/00346543046003407

Shavelson, R. J., & Bolus, R. (1982). Self-concept: The interplay of theory and methods. *Journal of Educational Psychology*, *74*(1), 3–17. https://doi.org/10.1037/0022-0663.74.1.3

Siegel Robertson, J. (2000). Is attribution training a worthwhile classroom intervention for K–12 students with learning difficulties? *Educational Psychology Review*, *12*(1), 111–134. https://doi.org/10.1023/A:1009089118008

Skaalvik, E. M., & Hagtvet, K. A. (1990). Academic achievement and self-concept: An analysis of causal predominance in a developmental perspective. *Journal of Personality and Social Psychology*, *58*(2), 292.

Skaalvik, E. M., & Rankin, R. J. (1992). Math and Verbal Achievement and Self-Concepts: Testing the Internal/External Frame of Reference Model. *The Journal of Early Adolescence*, *12*(3), 267–279. https://doi.org/10.1177/0272431692012003003

Skaalvik, E. M., & Skaalvik, S. (2002). Internal and External Frames of Reference for Academic Self-Concept. *Educational Psychologist*, *37*(4), 233–244. https://doi.org/10.1207/S15326985EP3704_3

Soest, T. von, 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*(4), 592–608. https://doi.org/10.1037/pspp0000060

Song, I.-s., & Hattie, J. (1985). Relationships between self-concept and achievement. *Journal of Research in Personality*, *19*(4), 365–372. https://doi.org/10.1016/0092-6566(85)90005-4

Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin*, *139*(1), 213–240. https://doi.org/10.1037/a0028931

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*(2), 325–338. https://doi.org/10.1037/a0035133

Steinberg, L. D. (2008). *Adolescence* (8th ed.). Boston: McGraw-Hill Higher Education.

Steinmayr, R., & Spinath, B. (2015). Intelligence as a potential moderator in the internal/external frame of reference model. An exploratory analysis. *Journal for Educational Research Online/Journal für Bildungsforschung Online*, *7*(1), 198–218.

Strickhouser, J. E., & Zell, E. (2015). Self-evaluative effects of dimensional and social comparison. *Journal of Experimental Social Psychology*, *59*, 60–66. https://doi.org/10.1016/j.jesp.2015.03.001

Sukariyah, M. B., & Assaad, G. (2015). The effect of attribution retraining on the academic achievement of high school students in mathematics. *Procedia - Social and Behavioral Sciences*, *177*, 345–351. https://doi.org/10.1016/j.sbspro.2015.02.356

Susperreguy, M. I., Davis-Kean, P. E., Duckworth, K., & Chen, M. (2017). Self-concept predicts academic achievement across levels of the achievement distribution: domain specificity for math and reading. *Child Development.* Advance online publication. https://doi.org/10.1111/cdev.12924

Tetzner, J., Becker, M., & Maaz, K. (2016). Development in multiple areas of life in adolescence: Interrelations between academic achievement, perceived peer acceptance, and self-esteem. *International Journal of Behavioral Development*, *41*(6), 704–713. https://doi.org/10.1177/0165025416664432

Trautwein, U. (2003). *Schule und Selbstwert: Entwicklungsverlauf, Bedeutung von Kontextfaktoren und Effekte auf die Verhaltensebene [School and self-esteem. Course of development, meaning of context factors, and effects on behavior]*. Zugl.: Berlin, Freie Univ., Diss., 2002 u.d.T.: Trautwein, Ulrich: Die Entwicklung und Bedeutung des Selbstwertgefühls Jugendlicher im schulischen Kontext. *Pädagogische Psychologie und Entwicklungspsychologie: Vol. 36*. Münster: Waxmann.

Trautwein, U., Köller, O., & Baumert, J. (2004). Des einen Freud', der anderen Leid? [One man's joy is another man's sorrow?]. *Zeitschrift für Pädagogische Psychologie*, *18*(1), 15–29. https://doi.org/10.1024/1010-0652.18.1.15

Trautwein, U., Lüdtke, O., Marsh, H. W., & Nagy, G. (2009). Within-school social comparison: How students perceive the standing of their class predicts academic self-concept. *Journal of Educational Psychology*, *101*(4), 853–866. https://doi.org/10.1037/a0016306

Trautwein, U., & Möller, J. (2016). Self-Concept: Determinants and consequences of academic self-concept in school contexts. In A. A. Lipnevich, F. Preckel, & R. D. Roberts (Eds.), *The Springer Series on Human Exceptionality. Psychosocial skills and school systems in the 21st century* (pp. 187–214). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-28606-8_8

Trucco, E. M., Wright, A. G. C., & Colder, C. R. (2014). Stability and change of social goals in adolescence. *Journal of personality*, *82*(5), 379–389. https://doi.org/10.1111/jopy.12069

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*(2), 381–390. https://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*(1), 205–220. https://doi.org/10.1037/0022-3514.84.1.205

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, *39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3

Van de Schoot, R., Schmidt, P., Beuckelaer, A. de, Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement Invariance. *Frontiers in psychology*, *6*, 1064. https://doi.org/10.3389/fpsyg.2015.01064

Van Zanden, B., Marsh, H. W., Seaton, M., Parker, P. D., Guo, J., & Duineveld, J. J. (2017). How well do parents know their adolescent children? Parent inferences of student self-concepts reflect dimensional comparison processes. *Learning and Instruction*, *47*, 25–32. https://doi.org/10.1016/j.learninstruc.2016.09.002

Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological methods*, *17*(2), 176–192. https://doi.org/10.1037/a0027543

Wagner, C., & Valtin, R. (2004). Determinants of self-esteem in young adolescents. In *Proceedings of the 3rd International Biennial Self Research Conference.* Retrieved from https://www.researchgate.net/profile/Renate_Valtin/publication/301700565_Determinants_of_self-esteem_in_young_adolescents/links/5723abf308aef9c00b811ac2.pdf

Wentzel, K. R., & Erdley, C. A. (1993). Strategies for making friends: Relations to social behavior and peer acceptance in early adolescence. *Developmental psychology*, *29*(5), 819–826. https://doi.org/10.1037/0012-1649.29.5.819

Wigfield, A., & Eccles, J. (2002). The development of competence beliefs, expectancies for success, and achievement values from Childhood through Adolescence. In A. Wigfield & J. Eccles (Eds.), *Development of Achievement Motivation* (pp. 91–120). Elsevier. https://doi.org/10.1016/b978-012750053-9/50006-1

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*(4), 552–565. https://doi.org/10.1037/0012-1649.27.4.552

Wilkinson, L. & Task Force on Statistical Inference, American Psychological Association, Science Directorate. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594–604.

Winne, P. H., & Marx, R. W. (1981, April). *Winne, P. H., & Marx, R. W. (1981, April). Convergent and discriminant validity in self-concept measurement.* Paper presented at the annual meeting of the American Educational Research Association, Los Angeles.

Wright, A. G. C., & Hallquist, M. N. (2014). Mixture modeling methods for the assessment of normal and abnormal personality, part II: Longitudinal models. *Journal of personality assessment*, *96*(3), 269–282. https://doi.org/10.1080/00223891.2013.830262

Young, J. F., & Mroczek, D. K. (2003). Predicting intraindividual self-concept trajectories during adolescence. *Journal of adolescence*, *26*(5), 586–600. https://doi.org/10.1016/S0140-1971(03)00058-7

Yu, C.-Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (Doctoral dissertation). University of California, Los Angeles. Retrieved from https://www.statmodel.com/download/Yudissertation.pdf

Appendix A

Table A 1. Zero-order correlations, means and standard deviations, and reliabilities (Cronbach’s alpha) for self-esteem (SE), academic self-concept (ASC), social self-concept of acceptance (SCAC), and social self-concept of assertion (SCAS)

		T1: SE	T2: SE	T3: SE	T4: SE	T5: SE	T6: SE	T1: ASC	T2: ACS	T3: ASC	T4: ASC	T5: ASC	T6: ASC	T1: SCAC	T2: SCAC	T3: SCAC	T4: SCAC	T5: SCAC	T6: SCAC	T1: SCAS	T2: SCAS	T3: SCAS	T4: SCAS	T5: SCAS	T6: SCAS
T1: SE	*r*	–																							
	n	883																							
T2: SE	*r*	.56**	–																						
	n	774	820																						
T3: SE	*r*	.47**	.56**	–																					
	n	776	719	836																					
T4: SE	*r*	.38**	.43**	.52**	–																				
	n	722	667	688	811																				
T5: SE	*r*	.27**	.32**	.36**	.46**	–																			
	n	625	584	595	592	813																			
T6:SE	*r*	.27**	.33**	.35**	.34**	.44**	–																		
	n	414	388	400	394	423	477																		
T1: ASC	*r*	.57**	.40**	.33**	.25**	.13**	.18**	–																	
	n	882	773	775	721	624	414	882																	
T2: ASC	*r*	.39**	.60**	.43**	.32**	.23**	.23**	.54**	–																
	n	774	820	719	667	584	388	773	820																
T3: ASC	*r*	.29**	.39**	.58**	.40**	.29**	.21**	.47**	.59**	–															
	n	775	718	834	687	593	399	774	718	835															
T4: ASC	*r*	.23**	.34**	.36**	.47**	.25**	.17**	.39**	.48**	.56**	–														
	n	722	667	688	811	592	394	721	667	687	811														
T5: ASC	*r*	.19**	.27**	.29**	.22**	.37**	.19**	.33*	.37**	.48**	.52**	–													
	n	625	584	595	592	813	423	624	584	593	592	813													
T6: ASC	*r*	.12*	.17**	.14**	.10*	.18**	.33**	.32**	.31**	.34**	.42**	.52**	–												

	n	414	388	400	394	423	477	414	388	399	394	423	477												
T1: SCAC	*r*	.30**	.21**	.18**	.22**	.10*	.15**	.27**	.17**	.20**	.14**	.08*	.05	–											
	n	883	774	776	722	625	414	882	774	775	722	625	414	883											
T2: SCAC	*r*	.19**	.26**	.24**	.20**	.17**	.12*	.19**	.21**	.20**	.15**	.12**	.04	.46**	–										
	n	775	820	720	668	585	388	774	820	719	668	585	388	775	821										
T3: SCAC	*r*	.23**	.21**	.34**	.23**	.21**	.21**	.20**	.24**	.25**	.16**	.15**	.11*	.42**	.57**	–									
	n	776	719	836	688	595	400	775	719	834	688	595	400	776	720	836									
T4: SCAC	*r*	.10**	.11**	.20**	.37**	.14**	.15**	.12**	.12**	.15**	.20**	.06	.08	.33**	.33**	.45**	–								
	n	722	667	688	811	592	394	721	667	687	811	592	394	722	668	688	811								
T5: SCAC	*r*	-.01	.09*	.07	.14**	.31**	.18**	-.03	.07	.12**	.08*	.07	.03	.26**	.23**	.27**	.34**	–							
	n	625	584	595	592	813	423	624	584	593	592	813	423	625	585	595	592	813							
T6: SCAC	*r*	.11*	.16**	.08	.17**	.19**	.35**	.10	.11*	.05	.09	-.03	.05	.23**	.15**	.20**	.35**	.39**	–						
	n	414	388	400	394	423	477	414	388	399	394	423	477	414	388	400	394	423	477						
T1: SCAS	*r*	.30**	.24**	.21**	.24**	.17**	.12*	.23**	.19**	.15**	.18**	.15**	.12*	.36**	.28**	.24**	.28**	.15**	.06	–					
	n	883	774	776	722	625	414	882	774	775	722	625	414	883	775	776	722	625	414	883					
T2: SCAS	*r*	.18**	.28**	.24**	.20**	.17**	.08	.18**	.24**	.18**	.19**	.12**	.08	.25**	.45**	.34**	.25**	.15**	.09	.54**	–				
	n	774	820	719	667	584	388	773	820	718	667	584	388	774	820	719	667	584	388	774	820				
T3: SCAS	*r*	.23**	.21**	.28**	.25**	.26**	.14**	.17**	.19**	.19**	.23**	.20**	.13*	.26**	.33**	.43**	.30**	.21**	.09	.51**	.59**	–			
	n	776	719	836	688	595	400	775	719	834	688	595	400	776	720	836	688	595	400	776	719	836			
T4: SCAS	*r*	.14**	.15**	.22**	.32**	.20**	.17**	.14**	.13**	.14**	.20**	.13**	.14**	.25**	.30**	.31**	.43**	.23**	.18**	.43**	.46**	.55**	–		
	n	722	667	688	811	592	394	721	667	687	811	592	394	722	668	688	811	592	394	722	667	688	811		
T5: SCAS	*r*	.06	.16**	.16**	.16**	.35**	.22**	.07	.12**	.13**	.08*	.13**	.05	.13**	.17**	.19**	.18**	.46**	.21**	.29**	.32**	.38**	.42**	–	
	n	625	584	595	592	813	423	624	584	593	592	813	423	625	585	595	592	813	423	625	584	595	592	813	
T6: SCAS	*r*	.12*	.23**	.16**	.21**	.14**	.33**	.11*	.17**	.12*	.12*	.03	.14**	.08	.12*	.14**	.18**	.18**	.32**	.29**	.24**	.34**	.36**	.44**	–
	n	414	388	400	394	423	477	414	388	399	394	423	477	414	388	400	394	423	477	414	388	400	394	423	477
	M	4.14	4.09	3.98	3.99	4.01	3.94	4.26	4.04	3.92	3.79	3.60	3.60	4.26	4.37	4.25	4.35	4.42	4.57	3.60	3.67	3.57	3.79	3.92	3.88
	SD	.62	.67	.70	.72	.67	.65	.64	.68	.76	.79	.83	.77	.83	.82	.89	.88	.82	.59	.95	1.01	1.02	.94	.87	.81
	α	.69	.69	.64	.70	.64	.73	.75	.77	.82	.81	.83	.84	.71	.77	.81	.82	.84	.80	.63	.71	.71	.71	.75	.79

Note. * $p < .05$. ** $p < .01$. *** $p < .001$. T1 = Time 1; T2 = Time 2; T3 = Time 3; T4 = Time 4; T5 = Time 5; T6 = Time 6.

Table A 2. Measurement invariance testing of self-esteem (SE), academic self-concept (ASC), social self-concept of acceptance (SCAC), and social self-concept of assertion (SCAS) across six time points

Construct	Invariance	*χ2 (df), SCF*	*p*	CFI	RMSEA	SRMR	*Δ* CFI
SE	Configural	236.415(177), 1.2145	.002	.986	.017 [.011, .022]	.037	
	Metric	264.872 (192), 1.2082	<.001	.983	.018 [.012, .023]	.044	.003
	Scalar	394.198 (207), 1.2021	<.001	.957	.028 [.024, .032]	.049	.029
	Partial Scalar [GCON4_1 free]	357.755 (206), 1.2031	<.001	.965	.025 [.021, .029]	.044	.018
	Partial Scalar [GCON4_1, GCON2_9 free]	338.010 (205), 1.1983	<.001	.970	.024 [.019, .028]	.045	.013
	Partial Scalar [GCON4_1, GCON2_9, GCON4_9 free]	317.492 (204), 1.1981	<.001	.974	.022 [.017, .026]	.045	.009
ASC	Configural	116.752 (75), 1.2239	.002	.991	.022 [.014, .029]	.032	
	Metric	132.497 (85), 1.2247	<.001	.990	.022 [.014, .029]	.048	.001
	Scalar	170.685 (95), 1.2247	<.001	.984	.026 [.020, .032]	.054	.006
SCAC	Configural	70.066 (75), 1.2411	.640	1	.000 [.000, .014]	.036	
	Metric	85.209 (85), 1.2994	.473	1	.001 [.000, .016]	.039	.000
	Scalar	150.467 (95), 1.2820	<.001	.983	.022 [.015, .029]	.047	.017
	Partial Scalar [SSAN1_1 free]	117.715 (94), 1.2812	.050	.993	.015 [.001, .023]	.042	.007
SCAS	Configural	96.871 (75), 1.1021	.046	.994	.016 [.002, .024]	.028	
	Metric	123.984 (85), 1.0896	.004	.989	.020 [.012, .027]	.040	.005

Scalar	160.306 (95), 1.0931	<.001	.981	.024 [.018, .031]	.045	.008

Note. *N* = 1,163. SCF = Scaling Correction Factor; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; SRMR = Standardized Root Mean Square Residual. Item: GCON4 “I am able to do things as well as most other people”; item: GCON2 “I feel that I have a number of good qualities”; item: SSAN1 “In class, I sometimes feel like an outsider”.

Table A 3. Model-based concurrent and cross-time correlations between self-esteem (SE), academic self-concept (ASC), social self-concept of acceptance (SCAC), and social self-concept of assertion (SCAS) across six time-points

	T1: SE	T2: SE	T3: SE	T4: SE	T5: SE	T6: SE	T1: ASC	T2: ACS	T3: ASC	T4: ASC	T5: ASC	T6: ASC	T1: SCA	T2: SCA	T3: SCA	T4: SCA	T5: SCA	T6: SCA	T1: SCA	T2: SCA	T3: SCA	T4: SCA	T5: SCA	T6: SCA
T1: SE	-																							
T2: SE	.73***	-																						
T3: SE	.55***	.75***	-																					
T4: SE	.39***	.53***	.69***	-																				
T5: SE	.26***	.34***	.44***	.62***	-																			
T6: SE	.15***	.20***	.25***	.36***	.57***	-																		
T1: ASC	.72***	.50***	.37***	.25***	.15***	.06*	-																	
T2:ASC	.49***	.77***	.55***	.37***	.22***	.09**	.67***	-																
T3: ASC	.31***	.51***	.71***	.46***	.26***	.10*	.50***	.73***	-															
T4: ASC	.22***	.36***	.48***	.58***	.32***	.10^{t}	.37***	.53***	.72***	-														
T5: ASC	.11**	.20***	.26***	.27***	.47***	.16**	.25***	.35***	.47***	.66***	-													
T6: ASC	.04	.10**	.13**	.11*	.20***	.32***	.16***	.22***	.30***	.42***	.64***	-												
T1: SCAC	.45***	.28***	.26***	.18***	.11***	.05*	.37***	.24***	.21***	.15***	.09**	.06*	-											
T2: SCAC	.28***	.34***	.33***	.24***	.14***	.06*	.26***	.26***	.24***	.18***	.11**	.06*	.64***	-										
T3: SCAC	.24***	.30***	.47***	.32***	.18***	.09**	.23***	.25***	.30***	.20***	.10**	.05†	.46***	.72***	-									
T4: SCAC	.11*	.14*	.22***	.46***	.23***	.10*	.14***	.14**	.16***	.24***	.09^{t}	.04	.26***	.41***	.56***	-								

T5: SCAC	.05†	.06*	.09*	.18**	.42***	.19*	.08**	.08*	.09*	.14*	.09*	.02	.13***	.19***	.26***	.44***	-							
T6: SCAC	.03	.03	.05^{t}	.11*	.23***	.42***	.01	.00	-.00	-.01	-.06	.02	.05**	.08**	.12***	.21***	.47***	-						
T1: SCAS	.45***	.37***	.31***	.28***	.21***	.11**	.34***	.30***	.23***	.23***	.17***	.11**	.53***	.42***	.34***	.23***	.13***	.05†	-					
T2: SCAS	.29***	.39***	.35***	.32***	.24***	.12*	.27***	.30***	.24***	.26***	.20***	.12***	.42***	.61***	.48***	.32***	.18***	.07†	.78***	-				
T3: SCAS	.26***	.34***	.39***	.37***	.29***	.15**	.22***	.26***	.24***	.28***	.22***	.14**	.32***	.47***	.56***	.39***	.22***	.08*	.63***	.81***	-			
T4: SCAS	.20***	.27***	.31***	.41***	.33***	.18**	.17***	.20***	.19***	.22***	.19***	.12*	.24***	.35***	.41***	.53***	.29***	.12*	.49***	.63***	.78***	-		
T5: SCAS	.13***	.18***	.21***	.26***	.50***	.28***	.11***	.14***	.14**	.16**	.17**	.10	.13***	.18***	.21***	.23***	.57***	.25***	.29***	.36***	.45***	.57***	-	
T6: SCAS	.07***	.09**	.10**	.12*	.22**	.40***	.04†	.05†	.04	.03	.02	.15**	.06**	.08**	.09**	.08†	.23**	.40***	.15***	.19***	.24***	.31***	.55***	-

Note. Results from the tech 4 Output in Mplus regarding Model 1. † $p < .10$ * $p < .05$. ** $p < .01$. *** $p < .001$. T1 = Time 1; T2 = Time 2; T3 = Time 3; T4 = Time 4; T5 = Time 5; T6 = Time 6.

Table A 4. Results of the second-order latent growth curve analyses with linear slopes for self-esteem (SE), academic self-concept (ASC), social self-concept of acceptance (SCAC), and social self-concept of assertion (SCAS) across six time points

		SE	ASC	SCAC	SCAS
Model A: linear growth					
Model fit					
	$\chi^2(df)$, p, Scaling correction factor	396.368 (220), p < .001, 1.1990	437.313 (111), p < .001, 1.2426	173.176 (110), p < .001, 1.3269	246.282 (111), p < .001, 1.1132
	CFI	.960	.931	.981	.961
	RMSEA	.026 [.022,.030]	.050 [.045, .055]	.022 [.016, .028]	0.032 [.027, .038]
	SRMR	.061	.142	.062	.063

Parameters		Utandardized Estimates	SE	Utandardized Estimates	SE	Utandardized Estimates	SE	Utandardized Estimates	SE
	Means								
	Intercept	4.046***	.021	4.048***	.023	4.297***	.028	3.625***	.032
	Slope linear	-.012	.007	-.110***	.006	.038***	.006	.052***	.006
	Variances								
	Intercept	.211***	.013	.228***	.016	.339***	.032	.491***	.036
	Slope linear	.006***	.001	.009***	.001	.010***	.002	.013***	.002
	Covariance slope x intercept	.015***	.003	-.004	.004	-.037***	.005	-.050***	.006
Model B: quadratic growth									
Model fit									
	χ^2 (*df*), p SCF	360.369 (216), *p* < .001, 1.1958		280.809 (107), *p* < .001, 1.2507		165.884 (106), *p* < .001, 1.3198		215.357(107), *p* < .001, 1.1074	
	$\Delta\ \chi^2$ (*df*), p	8.706 (4), *p* = .069		187.341 (4), *p* < .001		7.164 (4), *p* = .127		28.127 (4), *p* < .001	
	CFI	.967		.963		.982		.969	
	RMSEA	.024 [.020, .028]		.037 [.032,.043]		.022 [.015, .028]		.030 [.024, .035]	
	SRMR	.056		.089		.059		.058	
Parameters		Unstandardized Estimates	SE	Unstandardized Estimates	SE	Unstandardized Estimates	SE	Unstandardized Estimates	SE
	Means								
	Intercept	4.070***	.020	4.135***	.023	4.317***	.029	3.572***	.029
	Slope linear	-.057**	.020	-.284***	.023	.004	.020	.141***	.018
	Slope quadratic	.008*	.003	.029***	.003	.005t	.003	-.014***	.003
	Variances								

Intercept	.224***	.014	.206***	.018	.342***	.037	.514***	.037
Slope linear	.060***	.013	.072***	.015	.049*	.025	.063**	.022
Slope quadratic	.001*	.000	.002**	.000	.001	.001	.001	.001
Covariances								
Intercept x Slope linear	-.039***	.011	.009	.011	-.049*	.023	-.083***	.021
Intercept x Slope quadratic	.003*	.002	.003	.002	.002	.003	.005	.003
Slope linear x Slope quadratic	-.007**	.002	-.010***	.003	-.005	.004	-.007	.004

Note. SCF = Scaling Correction Factor; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; SRMR = Standardized Root Mean Square Residual.* $p < .05$. ** $p < .01$. *** $p < .001$.

Table A 5. Model 1: Standardized parameter estimates for the cross-lagged paths and correlations between residual variances regarding self-esteem (SE), academic self-concept (ASC), social self-concept of acceptance (SCAC), social self-concept of assertion (SCAS), across the six time points (t1-t6)

Stability	SE	ASC	SCAC	SCAS
t1-t2	.756, $p < .001$	.668, $p < .001$	.576, $p < .001$	.797, $p < .001$
t2-t3	.756, $p < .001$	.842, $p < .001$	.670, $p < .001$	.825, $p < .001$
t3-t4	.695, $p < .001$	.762, $p < .001$	.521, $p < .001$	.795, $p < .001$
t4-t5	.653, $p <. 001$	.754, $p < .001$	.420, $p < .001$	.602, $p < .001$
t5-t6	.669, $p < .001$	.708, $p < .001$	.460, $p < .001$	.632, $p < .001$
Bottom-up effects				
	ASC → SE	SCAC → SE	SCAS → SE	

t1-t2	-.039, *p* = .705	-.091, *p* = .204	.089, *p* = .140	
t2-t3	-.058, *p* = .550	.074, *p* = .323	.029, *p* = .707	
t3-t4	-.046, *p* = .519	-.077, *p* = .206	.155, *p* = .008	
t4-t5	-.060, *p* = .383	-.138, *p* = .055	.146, *p* = .029	
t5-t6	-.155, *p* =.026	-.092, *p* = .387	.026, *p* = .789	
Top-down effects				
	SE → ASC	SE → SCAC	SE → SCAS	
t1-t2	-.015, *p* = .863	-.073, *p* = .320	-.138, *p* = .060	
t2-t3	-.170, *p* = .017	.026, *p* = .728	.053, *p* = .608	
t3-t4	-.088, *p* = .296	-.124, *p* = .130	.015, *p* = .851	
t4-t5	-.188, *p* = .016	-.089, *p* = .246	.054, *p* = .465	
t5-t6	-.158, *p* = .018	.135, *p* =. 113	-.007, *p* = .946	
Tradimeional effects				
	SCAC →ASC	ASC→ SCAC	ASC→ SCAS	SCAS →ASC
t1-t2	-.067, *p* = .195	.060, *p* = .465	.094, *p* = .231	.120, *p* = .003
t2-t3	.088, *p* = .098	.045, *p* = .473	-.024, *p*= .769	-.004, *p* = .936
t3-t4	-.092, *p* = .159	.065, *p* = .330	.006, *p* = .931	.187, *p* = .004
t4-t5	-.083, *p* =. 032	.067, *p* = .392	.019, *p* = .779	.143, *p* = .005
t5-t6	-.012, *p* = .888	-.161, *p* = .016	-.075, *p* = .339	.063, *p* = .429
	SCAC→ SCAS	SCAS→ SCAC		
t1-t2	.017, *p* =.814	.129, *p* = .022		
t2-t3	-.048, *p* =.543	.129, *p* = .054		

t3-t4	-.043, p =.485	.092, p = .142		
t4-t5	-.116, p =.084	.092, p = .142		
t5-t6	-.124, p =.196	-.049, p = .590		
Correlations				
	ASC ↔ SE	ASC ↔ SCAC	ASC ↔ SCAS	SCAC ↔ SE
t1	.720, p < .001	.371, p <. 001	.335, p < .001	.446, p < .001
t2	.847, p < .001	.133, p =. 066	.139, p = .121	.299, p < .001
t3	.776, p < .001	.174, p = .005	.115, p = .182	.403, p < .001
t4	.511, p < .001	.200, p = .008	-.010, p = .916	.510, p < .001
t5	.564, p < .001	.029, p = .708	.066, p = .442	.476, p < .001
t6	.477, p < .001	.139, p = .088	.227, p < .001	.426, p < .001
	SCAS ↔ SE	SCAC ↔ SCAS		
t1	.448, p < .001	.534, p < .001		
t2	.344, p <. 001	.587, p < .001		
t3	.241, p < .001	.476, p < .001		
t4	.261, p < .001	.462, p < .001		
t5	.451, p < .001	.595, p < .001		
t6	.398, p < .001	.412, p < .001		

Note. T1 = Time 1; T2 = Time 2; T3 = Time 3; T4 = Time 4; T5 = Time 5; T6 = Time 6.

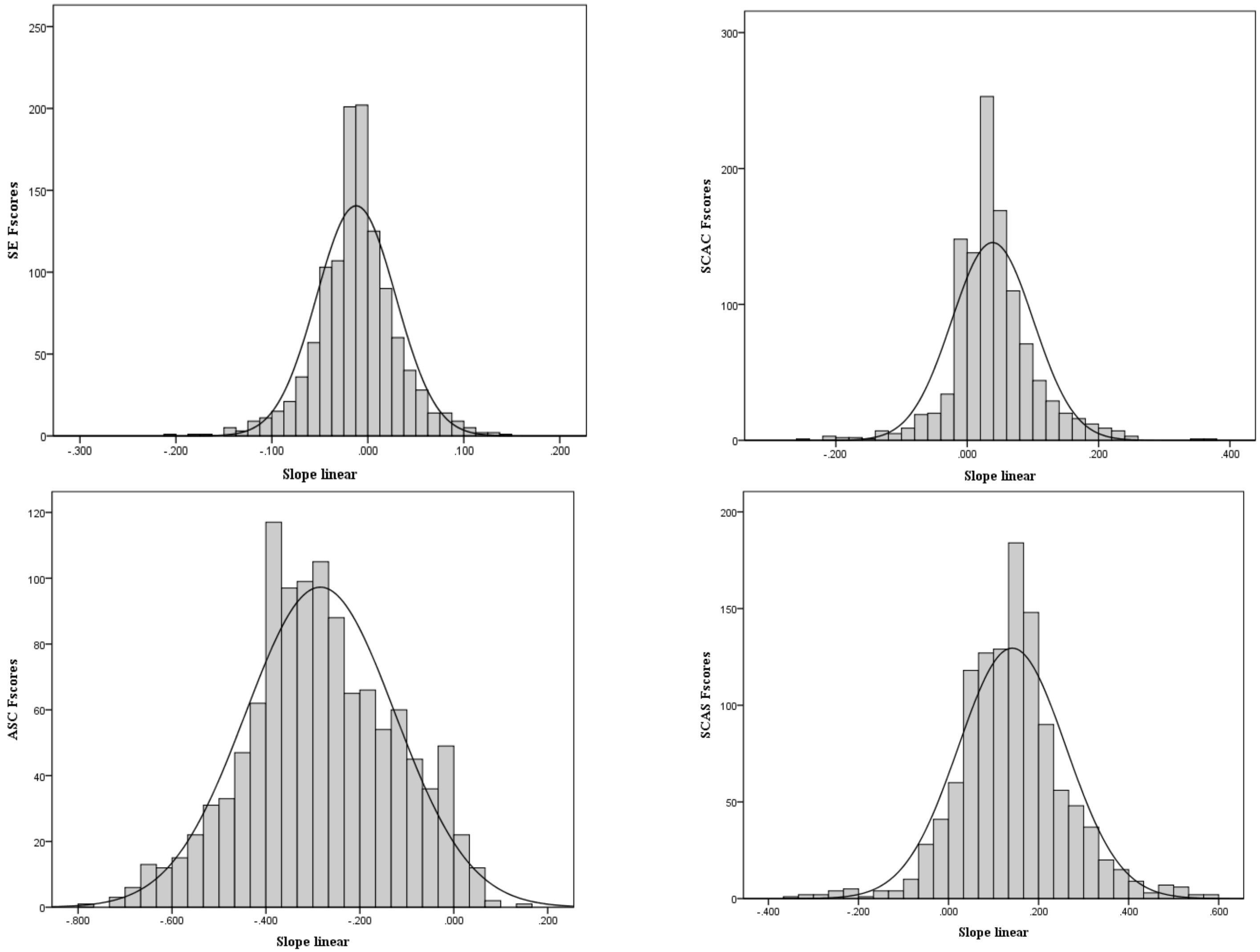


Figure A 1. Histograms for the individual estimated linear slope factor from second-order latent growth curve model for self-esteem (SE), general academic self-concept (ASC), social self-concept of acceptance (SCAC), and social self-concept of assertion (SCAS)

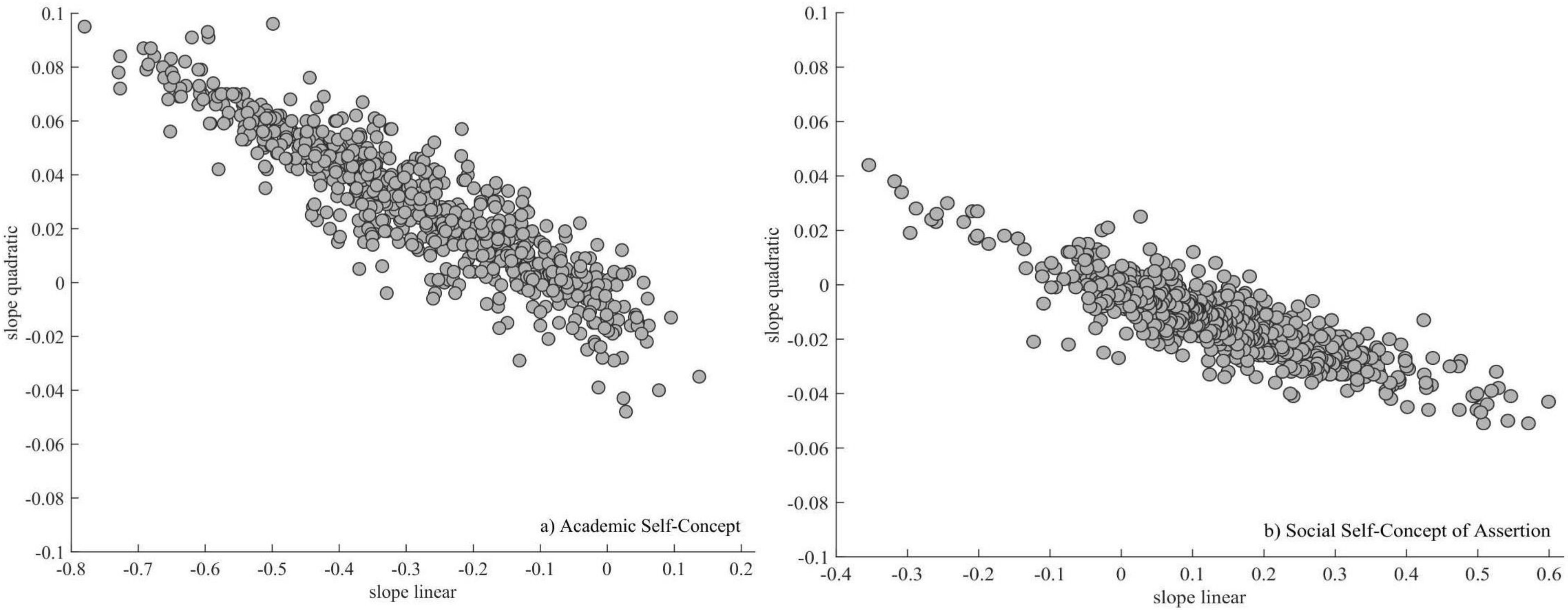


Figure A 2. Bivariate scatterplots for linear and quadratic latent growth curve models of (a) general academic self-concept, and (b) social self-concept of assertion

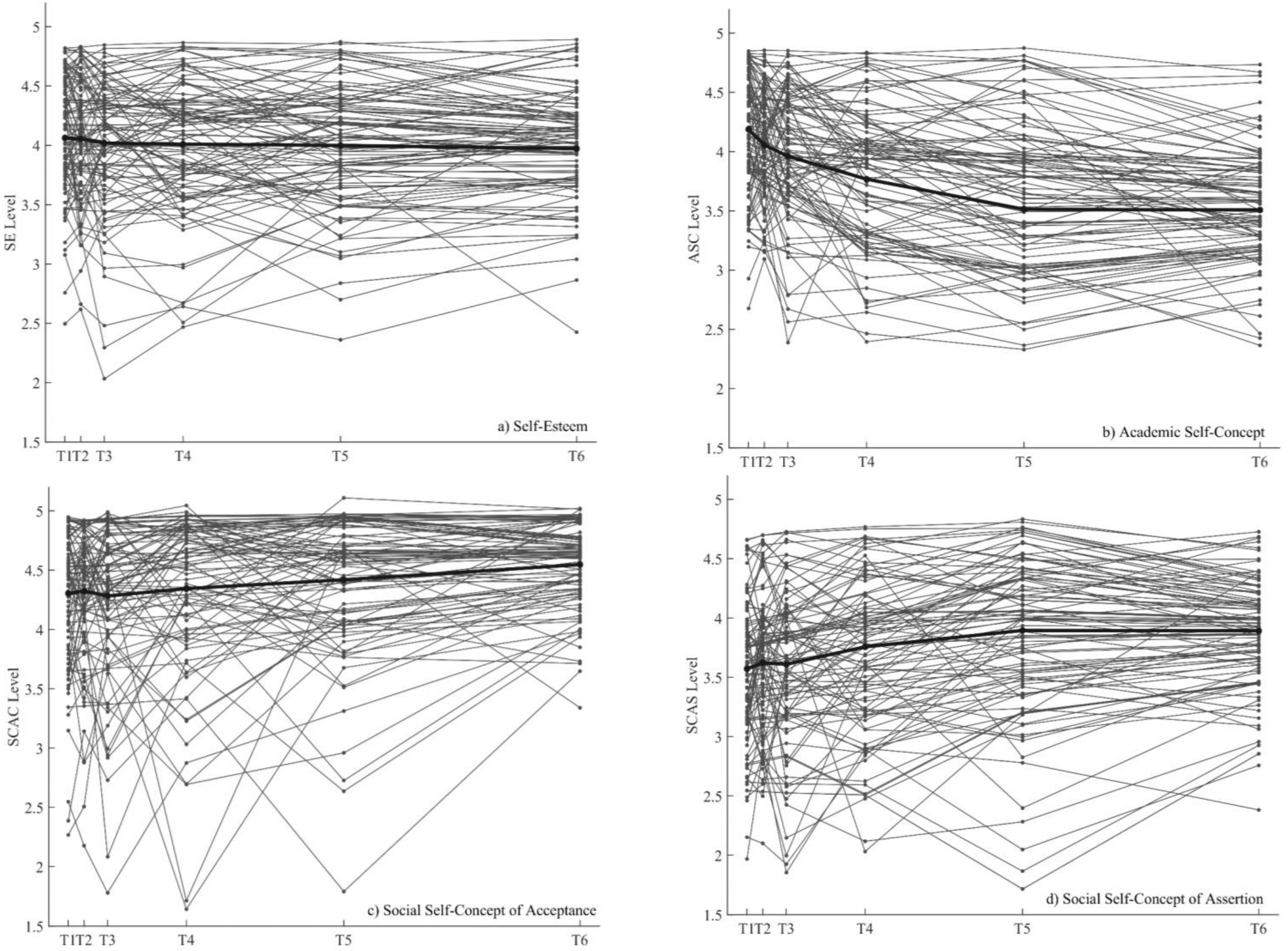


Figure A 3. Individual-data plots (spaghetti plots) of (a) self-esteem, (b) general academic self-concept, (c) social self-concept of acceptance, and (d) social self-concept of assertion (N = 100 randomly selected individuals). T1 = Time 1; T2 = Time 2; T3 = Time 3; T4 = Time 4; T5 = Time 5; T6 = Time 6.

Appendix B

Table B 1. Standardized factor-loadings for the NMS model (M1: NMS_3sf) and the first-order correlated factor model of academic self-concept (ASC) (M3: FOCF_4f), both with math ASC, writing ASC, and reading ASC factors, for grades 1 to 4

Grade	1		2		3		4	
Model	M1	M3	M1	M3	M1	M3	M1	M3
gASC								
a1	.733	.677	.790	.791	.742	.755	.737	.740
a2	.799	.805	.749	.762	.807	.862	.802	.804
a3	.839	.758	.796	.745	.875	.815	.865	.875
m1	.722		.675		.680		.673	
m2	.696		.636		.643		.661	
m3	.700		.651		.654		.673	
r1	.512		.483		.533		.478	
r2	.549		.510		.514		.530	
r3	.625		.546		.596		.608	
w1	.511		.558		.519		.492	
w2	.578		.568		.532		.546	
w3	.660		.639		.574		.584	
mASC								
m1	.592	.923	.624	.931	.649	.954	.640	.936
m2	.653	.941	.671	.900	.700	.929	.717	.958
m3	.468	.866	.490	.828	.508	.839	.521	.865
rASC								
r1	.755	.880	.789	.908	.808	.960	.797	.915
r2	.768	.949	.790	.937	.740	.892	.839	.996
r3	.540	.868	.523	.800	.530	.813	.594	.865
wASC								
w1	.656	.804	.733	.903	.780	.917	.795	.933
w2	.704	.876	.708	.913	.760	.917	.838	.996
w3	.509	.890	.504	.828	.518	.803	.566	.865

Note. All factor-loadings were significant ($p < .001$). a1/a2/a3 = Items of general ASC; m1/m2/m3 = Items of math ASC; r1/r2/r3 = Items of reading ASC; w1/w2/w3 = Items of writing ASC. Despite equal nomenclature in M1 and M3, mASC = math ASC, wASC = writing ASC and rASC = reading ASC are first-order correlated factors in the FOCF_4f model and specific factors that are residualized by gASC in the NMS_3sf model; thus, factors from both models are not directly comparable. M1: NMS_3sf = Nested Marsh/Shavelson-model of ASC with math ASC, writing ASC, and reading ASC as correlated specific factors; M3: FOCF_4f = First-order correlated factor model with general ASC, math ASC, writing ASC, and reading ASC as factors.

Table B 2. Reliability coefficient omega (ω; McDonald, 1999) for factors of general academic self-concept (ASC) and math ASC, writing ASC, and reading ASC by grade Level and model (M1: NMS_3sf; M3: FOCF_4f)

Grade	1		2		3		4	
Factors	M1	M3	M1	M3	M1	M3	M1	M3
gASC	.95	.81	.95	.81	.95	.85	.96	.85
mASC	.85	.93	.83	.92	.87	.92	.88	.95
rASC	.88	.98	.87	.92	.88	.92	.92	.94
wASC	.82	.89	.85	.92	.86	.92	.91	.94

Note. Despite equal nomenclature in M1 and M3, mASC = math ASC, wASC = writing ASC and rASC=reading ASC are first-order correlated factors in the FOCF_4f model and specific factors that are residualized by gASC in the NMS_3sf model; thus, factors from both models are not directly comparable. M1: NMS_3sf = Nested Marsh/Shavelson-model of ASC with math ASC, writing ASC, and reading ASC as correlated specific factors; M3: FOCF_4f = First-order correlated factor model with general ASC, math ASC, writing ASC, and reading ASC as factors.

Table B 3. Latent means in grades 1 to 4 of the factors derived from the NMS model (M1: NMS_3sf) and the first-order correlated factor model (M3: FOCF_4f), both with math ASC, writing ASC, and reading ASC

Grade		1	2	3	4
Model	Factors				
		Reference Group: Grade 1			
M1	gASC	0	-.43†	-1.09***	-1.29***
M3	gASC	0	-.36*	-.73***	-.77***
M3	mASC	0	-.65†	-1.72***	-1.74***
M3	rASC	0	-.39	-1.00***	-1.20***
M3	wASC	0	-.49*	-.98***	-1.24***

		Reference Group: Grade 2			
M1	gASC		0	-.59***	-.74***
M3	gASC		0	-.37***	-.41***
M3	mASC		0	-1.03***	-1.04***
M3	rASC		0	-.58*	-.78***
M3	wASC		0	-.50**	-.76***
		Reference Group: Grade 3			
M1	gASC			0	-.08
M3	gASC			0	-.06
M3	mASC			0	-.01
M3	rASC			0	-.20
M3	wASC			0	-.28

Note. Despite equal nomenclature in M1 and M3, mASC = math ASC, wASC = writing ASC, and rASC = reading ASC are first-order correlated factors in the FOCF_4f model and specific factors that are residualized by gASC in the NMS_3sf model; thus, factors from both models are not directly comparable. M1: NMS_3sf = Nested Marsh/Shavelson model of ASC with math, writing, and reading as correlated specific factors; M3: FOCF_4f = First-order correlated factor model with general ASC, math ASC, writing ASC, and reading ASC as factors. †p < .10 *p < .05 **p < .01 ***p < .001.

Latent mean level differences were estimated in a multigroup model of the scalar invariant model by setting the latent mean in the reference group (i.e., grade level) to zero and estimating the mean in the other group. The estimated latent mean corresponds directly to the mean level difference between grade levels; the significance level indicates whether there is a significant mean difference. Results are presented in Table B3. Table B3 shows that there were significant negative latent mean differences in general ASC between grades 1 and 2 in the FOCF_4f model (M = -.36, p = .015), and between grades 2 and 3 in both models (NMS_3sf model: M = -.59, p < .001; FOCF_4f model: M = -.37, p <.001). Latent mean differences for reading ASC, writing ASC, and math ASC within the FOCF_4f model revealed that writing ASC declined from grades 1 to 2 [M = -.49, p < .001], whereas math ASC and reading ASC decreased only from grades 2 to 3 [M = -.59, p = .019]. No further significant changes were found in math ASC, writing ASC, and reading ASC between grades 3 and 4. The pattern of results indicated a lower general ASC, math ASC, writing ASC, and reading ASC at higher grade levels within the first three years of elementary school.

Appendix C

Measurement invariance testing

We tested metric measurement invariance in a multi-group analysis by comparing a self-concept Model with constrained factor loadings set invariant between the groups with a model without any equality constraints between the groups. Full metric invariance was assumed for most of the samples as indicated by anon-significant chi-square difference test between the two models (see Tables C1 and C2). In samples S1/4, S2/6, and S3/7, one equality constraint concerning a factor loading had to be released, so that partial metric invariance is confirmed (e.g., Dimitrov, 2010).

Table C 1. Measurement invariance testing between the groups of students having equal grades in math and German (group: equal) and those having unequal grades (group: unequal)

Study/grade		*N*	*χ2 (df)*	CFI	RMSEA	SRMR	*Δ χ2 (Δ df)*	*p*
S 1/3	Group: unequal	180	35.061 (22)*, 1.1277	.977	.057 [.014,.092]	.062		
	Group: equal	146	43.723 (22)**, 1.2840	.942	.082 [.046,.118]	.063		
	Configural Model	326	79.426 (44)***, 1.2046	.962	.070 [.045,.095]	.063		
	Metric Model	326	89.877 (50)***, 1.1901	.958	.070 [.046,.093]	.074	10.414 (6)	.108
S 1/4	Group: unequal	506	54.647 (22)***, 1.1582	.982	.054 [.036,.072]	.038		
	Group: equal	302	33.435 (22)***, 1.2094	.986	.041 [.000,.068]	.043		
	Configural Model	808	87.624 (44)***, 1.1838	.983	.050 [.034,.065]	.040		
	Metric Model	808	102.456 (50)***,1.1921	.980	.051 [.037,.065]	.047		
	Partial Metric Model: MASC item 1 released[a]	808	95.755 (49)***, 1.1934	.982	.049 [.034,.063]	.045	8.2517 (5)	.143
S 2/5	Group: unequal	472	112.078 (32)***, 1.2980	.957	.073 [.058,.088]	.048		
	Group: equal	406	71.938 (32)***, 1.6553	.962	.055 [.038,.073]	.045		
	Configural Model	878	179.160 (64)***, 1.4767	.958	.065 [.055,.075]	.047		
	Metric Model	878	191.999 (72)***, 1.4630	.957	.062 [.051,.072]	.055	12.071 (8)	.150
S 2/6	Group: unequal	460	105.704 (32)***, 1.2671	.961	.071 [.056,.086]	.041		
	Group: equal	378	62.676 (32)***, 1.3402	.976	.050 [.031,.069]	.046		
	Configural Model	838	167.173 (64)***,1.3037	.968	.062 [.051,.074]	.045		
	Metric Model	838	184.890 (72)***,1.2660	.965	.061 [.050,.072]	.057	16.724 (8)	.170

	Partial Metric Model: MASC item 5 released[b]	838	178.347(71)***, 1.2729	.966	.060 [.049,.071]	.053	9.154 (7)	.242
S 2/8	Group: unequal	492	135.026 (32)***,1.3303	.953	.081 [.067,.095]	.039		
	Group: equal	349	120.446 (32)***,1.3592	.930	.089 [.072,.106]	.061		
	Configural Model	841	255.315 (64)***,1.3448	.944	.084 [.074,.095]	.049		
	Metric Model	841	267.949 (72)***,1.3157	.945	.080 [.070,.091]	.053	8.489 (8)	.387

Note. RMSEA=Root mean square error of approximation; CFI= Comparative fit index; SRMR=Standardized root mean square residual; SCF=scaling correction factor; df=degrees of freedom. MASC = Math self-concept; GASC = German self-concept. ****p*<.001 ***p*<.01 **p*<.05. [a]Item 1=”I am good at math.”, [b]Item 5= “I learn things quickly in math”.

Table C 2. Measurement invariance testing between the groups of students having equal grades in math and German (group: equal) and those having unequal grades (group: unequal)

Study/grade		*N*	*χ2 (df), SCF*	CFI	RMSEA	SRMR	*Δ χ2 (Δ df)*	*p*
S 3/5	Group: unequal	250	18.625 (12), 1.3103	.989	.047 [.000,.086]	.033		
	Group: equal	249	24.181 (12)*,1.1805	.972	.064 [.025,.101]	.046		
	Configural Model	499	42.516 (24)*,1.2454	.983	.056 [.026,.082]	.040		
	Metric Model	499	45.050 (29)*,1.2154	.985	.047 [.015,.073]	.045	1.685 (5)	.891
	Group: unequal	227	12.820 (12), 1.5134	.998	.017 [.000,.071]	.034		
S 3/6	Group: equal	234	23.276 (12)*,1.2790	.971	.063 [.022,.102]	.039		
	Configural Model	461	35.218 (24), 1.3963	.987	.045 [.000,.075]	.037		
	Metric Model	461	40.242 (29), 1.3224	.987	.041 [.000,.069]	.044	35.218 (5)	.525
S 3/7	Group: unequal	269	59.296 (13)***, 1.1084	.934	.121 [.091,.153]	.050		
	Group: equal	140	19.700 (12), 1.3392	.968	.068 [.000,.119]	.061		
	Configural Model	409	75.258 (24)***, 1.2239	.946	.102 [.077,.129]	.054		
	Metric Model	409	95.475 (29)***, 1.1619	.930	.106 [.083,.130]	.084	21.780 (5)	.001
	Partial Metric Model: GASC Item 2 released[a]	409	79.559 (28)***, 1.2080	.946	.095 [.071,.120]	.061	3.594 (4)	.464

Note. RMSEA=Root mean square error of approximation; CFI= Comparative fit index; SRMR=Standardized root mean square residual; SCF=Scaling correction factor; df=degrees of freedom. MASC = math self-concept; GASC = German self-concept. ***p<.001, and *p<.05. [a] Item 2= “I learn things quickly in German”.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe bzw. unerlaubte Hilfsmittel angefertigt, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Trier, den 2.05.2018