



The Structure of Academic Self-Concept: A Methodological Review and Empirical Illustration of Central Models

A. Katrin Arens 

DIPF/Leibniz Institute for Research and Information in Education

Malte Jansen

*Institute for Educational Quality Improvement and Centre
for International Student Assessment, Germany*

Franzis Preckel

University of Trier

Isabelle Schmidt

GESIS–Leibniz-Institute for the Social Sciences

Martin Brunner

University of Potsdam

The structure of academic self-concept (ASC) is assumed to be multidimensional and hierarchical. This methodological review considers the most central models depicting the structure of ASC: a higher-order factor model, the Marsh/Shavelson model, the nested Marsh/Shavelson model, a bifactor representation based on exploratory structural equation modeling, and a first-order factor model. We elaborate on how these models represent the theoretical assumptions on the structure of ASC and outline their inherent psychometric properties. We analyzed these models using a data set of German 10th grade students (N = 1,232) including a wide range of domain-specific ASCs as well as general ASC. The correlations among ASCs and between ASCs and academic achievement varied depending on the structural model used. We conclude with discussing recommendations for research purposes and advantages and limitations of each ASC model. Our approach may also guide research on other affective or motivational constructs (e.g., academic anxiety or interest).

KEYWORDS: academic self-concept, structural models, hierarchical constructs, multidimensional constructs, methodological review

Academic self-concept (ASC) is defined as the mental representation of one's own academic abilities in general and in different academic domains (e.g., Brunner et al., 2010). Researchers have also used the terms *self-concept of ability* (Helmke & van Aken, 1995) and *perceived cognitive competence* (Harter, 1982) with similar definitions. ASC has been an important construct in education research for several decades, as it relates to desirable outcomes such as higher educational aspirations, better attainment, and more favorable learning behavior (Marsh, 2007; Marsh & Craven, 2006).

For the purpose of this article, it is useful to separate two lines of research (Brunner et al., 2010; Edwards & Bagozzi, 2000; Marsh & Craven, 2006; Shavelson et al., 1976): The first line of research—also known as “within-network analyses” (Byrne, 1984, 1996)—addresses the internal structure of ASC. The second line of research—also known as “between-network analyses” (Byrne, 1984, 1996)—addresses the nomological network of ASC by investigating relations between ASC and outcomes such as academic achievement. These two lines of research are intertwined, given that studying between-network relations requires an appropriate structural model of ASC. In other words, before examining the theoretical and practical significance of ASC by relating it to other variables, it is essential to clarify the underlying conceptual characteristics of ASC. Consequently, the question of how to best capture the structure of ASC has stimulated much theoretical deliberation and empirical research.

Prior reviews on the structure of ASC (Byrne, 1984; Marsh & Hattie, 1996; Shavelson et al., 1976) are outdated and limited in scope because they did not cover the many currently available and applied ASC models. In addition, the purpose of previous work on different models on the structure of ASC (Brunner et al., 2010; Marsh, 1990b; Morin et al., 2016) was to introduce new models of ASC that addressed the challenges of previous models, rather than providing a comprehensive review and systematic comparison of existing models. The goal of this methodological review is therefore to provide an in-depth discussion of five central ASC models that have been most often applied or have recently been developed in contemporary ASC research: the higher-order factor model, the Marsh/Shavelson model, the nested Marsh/Shavelson model, a bifactor representation based on exploratory structural equation modeling (ESEM), and the first-order factor model.

This article consists of two parts. In Part 1, we elaborate on the five models on the structure of ASC within the framework of within-network analyses to describe the internal structure and components of ASC as assumed by these models. We also elaborate on the five models on the ASC structure within the framework of between-network analyses to discuss how these models test relations to outcome variables (e.g., academic achievement). In doing so, we highlight important similarities and differences concerning theoretical, conceptual, methodological, and psychometric characteristics of the ASC models. In Part 2, we illustrate the application of the five ASC models for a large data set of secondary school students from Germany. We investigate the different ASC models and examine how the correlations between academic achievement and ASC vary depending on the structural model of ASC applied. The syntax codes for all statistical analyses are provided in the supplements (in the online version of the journal) to facilitate the

application of these models in future research. We conclude by discussing the empirical findings from a theoretical and methodological perspective. In addition, we give recommendations concerning which research questions can be best addressed with the different ASC models and outline the major shortcomings of each model. Given that other vital constructs in education research (e.g., academic anxiety, academic interest) share theoretical and psychometric underpinnings with ASC (e.g., domain specificity; Gogol et al., 2017), the present article may also be relevant for researchers focusing on such constructs.

Part 1: Within-Network and Between-Network Analyses on the Structure of ASC

The seminal review by Shavelson et al. (1976) marked the beginning of empirical and psychometric ASC research as it made several empirically testable assumptions about the structure of self-concept (SC). Shavelson et al. (p. 411) defined SC as “a person’s perception of himself [sic]” and assumed SC to be both multidimensional and hierarchical in nature. Multidimensionality means that SC consists of different domain-specific facets tapping several domains of an individual’s life and experiences. Hierarchy means that the domain-specific SC facets are located on different generality levels. General SC is assumed to be located on the highest and most general level of the hierarchy and separated into ASC and non-ASC on a subordinate level.

According to Shavelson et al. (1976), ASC itself has a multidimensional and hierarchical structure. That is, students form separate SCs related to different school subjects or academic domains and these domain-specific ASCs can be combined into a general ASC (Marsh, 1990b).¹ Most of the ASC research has been conducted with students at school; thus, the academic connotation of ASC commonly refers to the school context (Marsh & Craven, 2006).

The model proposed by Shavelson et al. (1976) invoked a surge in research to empirically validate the assumptions of the multidimensional and hierarchical structure of ASC (i.e., within-network analyses; Byrne, 1996). The development and empirical evaluation of ASC models has benefitted from the evolution of structural equation modeling (SEM [e.g., Bollen, 1989; Kline, 2005]) that differentiates between manifest or observed variables (items) and latent, unobserved constructs (factors). The manifest variables operate as indicators for the latent constructs. Methodological advancements such as confirmatory factor analyses (CFAs) and their refinements have also contributed to the development of various ASC models (MacCallum & Austin, 2000). In the following, we describe the five central ASC models that are considered in detail in this article: the higher-order factor model, the Marsh/Shavelson model (Marsh, 1990b), the nested Marsh/Shavelson model (Brunner et al., 2010), a bifactor model implemented in ESEM (Asparouhov & Muthén, 2009), and the first-order factor model.

Within-Network Analyses: Models on the Structure of ASC

Higher-order Factor Model

The assumptions by Shavelson et al. (1976) implicated a higher-order factor model of ASC (Figure 1a). In this model, all domain-specific ASCs (e.g., verbal, math) formed first-order factors that load on a single higher-order factor

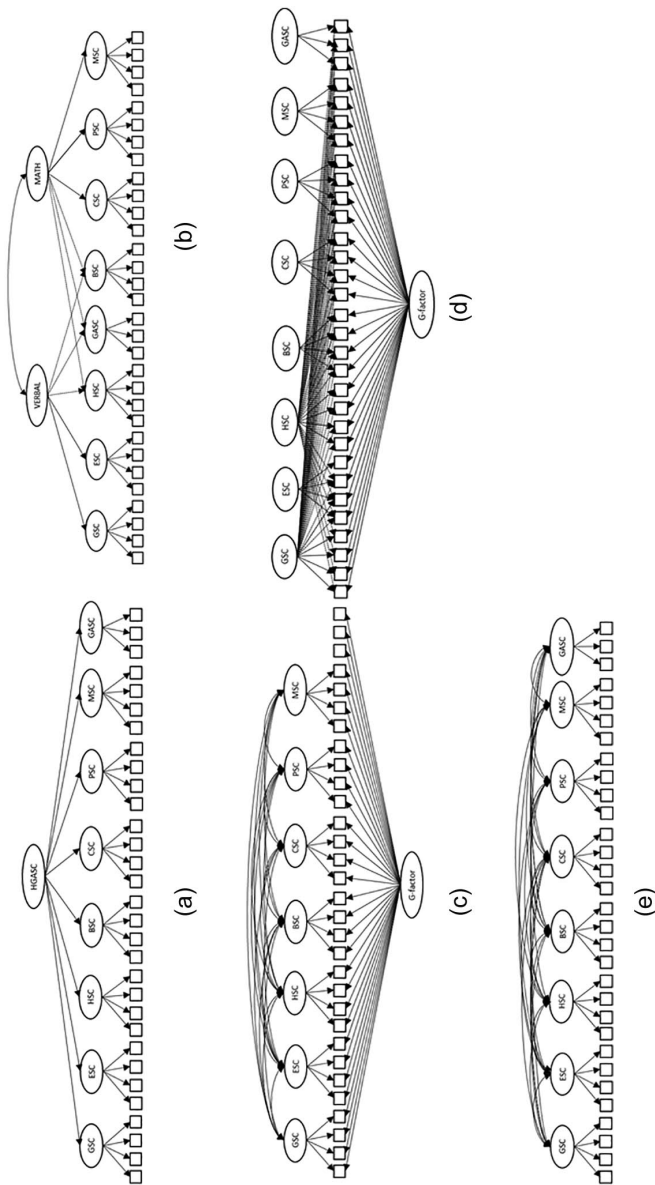


FIGURE 1 Structural models of academic self-concept: (a) Higher-order factor model; (b) Marsh/Shavelson model; (c) Nested Marsh/Shavelson model; (d) Bifactor-ESEM representation; and (e) First-order factor model. Residual terms as well as residual correlations between items using the same wordings were omitted for reasons of clarity. The three items in the nested Marsh/Shavelson model that load only on the G-factor refer to GASC. In the bifactor-ESEM model, only cross-loadings for GSC and HSC are illustrated to reduce complexity of the figure. Dashed lines represent cross-loadings. ESEM = exploratory structural equation modeling; HGASC = higher-order factor of general academic self-concept; MATH = higher-order math self-concept; VERBAL = higher-order verbal self-concept; GSC = German self-concept; ESC = English self-concept; HSC = history self-concept; BSC = biology self-concept; CSC = chemistry self-concept; PSC = physics self-concept; MSC = math self-concept; GASC = general academic self-concept.

representing general ASC (i.e., students' ASC across all domains). It is important to differentiate between two conceptualizations of general ASC. First, general ASC can be directly measured by items that assess how students perceive their abilities related to school in general and that are not tied to a specific domain or school subject (e.g., "I have always been good at school"). These items are then used to directly form general ASC either by deriving a latent first-order factor or by building a manifest scale score. Alternatively, general ASC can be extracted and aggregated from subordinate (first-order) factors of domain-specific ASCs. In this case, general ASC constitutes a higher-order factor that depicts the apex of a hierarchy of domain-specific ASCs. Such a higher-order factor of general ASC can either include or exclude a first-order general ASC factor (Yeung et al., 2000).² In the following, we refer to "GASC" for a first-order factor of general ASC that is directly measured by items capturing students' perceptions of their school-related abilities in general. We use the term "HGASC" to refer to a higher-order factor of general ASC composed of different first-order (domain-specific and GASC) factors.

The higher-order factor model clearly depicts the theoretical assumptions of Shavelson et al. (1976), but it often shows an inferior fit to the data compared to the alternative ASC models described below (Brunner et al., 2010), or even a poor fit (Marsh, 1987, 1990b). This is because the first-order math and verbal ASCs are consistently found to be nearly uncorrelated (Marsh, 1990a; Möller et al., 2009, 2020). It is reasonable to assume a higher-order factor only when the subordinate domain-specific factors are substantially correlated. In this case, the higher-order factor can effectively explain the common variance of the domain-specific factors (Brunner et al., 2012).

Marsh/Shavelson Model

The Marsh/Shavelson model was developed in response to the observation of nearly uncorrelated math and verbal ASCs and thus to remedy the shortcomings of the higher-order factor model (Marsh, 1990b). The model (Figure 1b) replaces the HGASC with two higher-order factors that are assumed to be nearly uncorrelated—a higher-order math ASC and a higher-order verbal ASC. The domain-specific ASCs form first-order factors that are located on a continuum from a verbal endpoint to a math endpoint. The position of a certain ASC on this continuum is reflected by the loadings on the higher-order math ASC and the higher-order verbal ASC. Hence, domain-specific ASCs can be classified as being either "math-like" or "verbal-like." The ASC related to students' main language of instruction (e.g., German for students in the German educational system; English for students in the English educational system) is assumed to represent the verbal endpoint of the ASC continuum in its purest form, and it should therefore have the highest loadings on the higher-order verbal ASC. Other verbal-like ASCs, such as the ASCs related to students' first foreign language or second foreign language, are located near the verbal endpoint of the ASC continuum. They are assumed to also load on the higher-order verbal ASC, but less strongly than the ASC related to students' main language of instruction. At the other end of the continuum, math ASC is assumed to represent the math endpoint in its purest form, and therefore should have the highest loadings on the higher-order math

ASC. Math-like ASCs such as science-related ASCs (e.g., ASCs in physics and chemistry) are assumed to be located close to the math endpoint of the ASC continuum, and therefore to load on the higher-order math ASC as well, but to a lesser degree than math ASC.

In some cases, domain-specific ASCs cannot be straightforwardly classified into either being math-like or being verbal-like. For example, history ASC and biology ASC are each assumed to have both math-like and verbal-like aspects and are thus located in the center of the continuum. Therefore, history and biology ASCs are assumed to load on both higher-order factors (Marsh et al., 2017). Finally, GASC is specified as a first-order factor that loads on both higher-order ASC factors, in line with its conceptualization as students' ASC related to all school subjects.

Nested Marsh/Shavelson Model

The nested Marsh/Shavelson model (Brunner et al., 2010; Figure 1c) capitalizes on advanced CFA models (Eid et al., 2003, 2017) and was developed to simultaneously account for the hierarchy and multidimensionality of ASCs. It reverts to the assumption of a hierarchically superordinate ASC that was abandoned in the Marsh/Shavelson model. In the nested Marsh/Shavelson model, all indicators (i.e., items) of domain-specific ASCs form an HGASC represented by a general (G-) factor. In addition, the indicators of domain-specific ASCs form first-order factors representing domain-specific ASCs (S-factors) that are nested under the G-factor. Items measuring GASC do not build a separate first-order S-factor but load on the G-factor only. The G-factor is specified to be uncorrelated with all domain-specific ASC factors (i.e., the S-factors). The different S-factors for the domain-specific ASC factors are allowed to correlate.

The nested Marsh/Shavelson model was empirically validated with samples of German students from elementary school (Schmidt et al., 2017) and secondary school (Brunner et al., 2008), with secondary school students from Luxembourg (Brunner et al., 2010; Gogol et al., 2017), and with secondary school students from 26 different countries (Brunner et al., 2009).

Bifactor-ESEM Representation

All models presented so far rely on CFA, which typically builds on the independent cluster model (ICM; McDonald, 1985), according to which each manifest item loads only on one single target factor without allowing any cross-loadings on other factors. However, the ICM approach might be too restrictive for multidimensional constructs such as ASC. Multidimensional constructs consist of different facets with some conceptual overlap, making cross-loadings theoretically plausible (e.g., Howard et al., 2018; Litalien et al., 2017; Marsh et al., 2009, 2011).

ESEM (Asparouhov & Muthén, 2009; Morin et al., 2013) has been recently established as a methodological framework within SEM that allows cross-loadings. When applying target rotation in ESEM, it is possible to a priori specify whether an item should have a main loading or a cross-loading on a specific factor. ESEM also allows the implementation of bifactor models where each manifest variable loads on a G-factor as well as on one or more S-factors (Morin et al.,

2016; Reise, 2012; Reise et al., 2011). Bifactor models are usually orthogonal models, meaning all S-factors are specified to be mutually uncorrelated and the correlations between the S-factors and the G-factor are also set to zero. Bifactor-ESEM representations have been successfully applied to model the joint structure of SC including academic and nonacademic facets (Arens & Morin, 2017; Morin et al., 2016). Still, bifactor-ESEM representations have not yet been applied to ASC only, although a bifactor-ESEM representation of ASC can well account for the hierarchy and multidimensionality of ASC as proposed by Shavelson et al. (1976). In a bifactor-ESEM representation of ASC, each ASC item has target loadings on a G-factor representing general ASC and on its corresponding S-factor. In addition, each ASC item has nontarget cross-loadings on the other S-factors (Figure 1d). The nested Marsh/Shavelson model can be conceptualized as an incomplete bifactor model because it does not include an S-factor for GASC items (Eid et al., 2017). The bifactor-ESEM representation, on the other hand, is a complete bifactor model because it comprises S-factors for all items including GASC items.

The First-Order Factor Model

A first-order factor model also builds on the ICM/CFA approach and assumes separate ASC factors for different domains or school subjects as well as for GASC (Figure 1e). The different ASC factors are allowed to correlate. However, the different ASC factors (including domain-specific ASCs as well as GASC) are not hierarchically related, instead all being located on the same level of hierarchy. Hence, the first-order factor model reflects the multidimensionality, but not the hierarchy, of ASC.

Between-Network Analyses: Outcome Relations of ASC

The importance of ASC in education research originates from its relations with outcome variables that are examined in the context of between-network analyses (Byrne, 1996). Academic achievement is probably the most frequently examined outcome variable of ASC (Hansford & Hattie, 1982; Marsh & Craven, 2006; Valentine et al., 2004). Academic achievement can be measured through standardized achievement test scores and school grades. Many empirical studies have found ASC to be more highly related to school grades than to standardized achievement test scores (Arens et al., 2017; Marsh et al., 2005; Möller et al., 2009, 2020). School grades are highly salient as they are directly and regularly communicated to the students (Marsh et al., 2014). Furthermore, students can easily compare their own school grades in one subject with other students' school grades in the same subject (social comparisons), with their own school grades in other subjects (dimensional comparisons), and with their own school grades in the same subject at previous points in time (temporal comparisons). These comparison processes have been conceptualized as an important mechanism of ASC formation (Möller et al., 2009, 2020; Wolff et al., 2019). The findings from between-network analyses have supported the domain specificity of ASC. Stronger relations have been found between ASC facets and achievement in matching domains (e.g., math ASC and math achievement) than between ASC facets and achievement in nonmatching domains (Marsh & Craven, 2006; Marsh

et al., 2017). Domain-specific ASCs are better able to explain and predict domain-specific achievement than indicators of general ASC such as GASC (Swann et al., 2007; Valentine et al., 2004).

Comparing Key Structural Characteristics Across ASC Models

In the following section, we discuss and compare the previously introduced models of ASC with respect to their assumptions about the nature and structure of ASC (within-network analyses) and with respect to the question of how achievement relations can be tested (between-network analyses). Table 1 summarizes the main aspects.

Multidimensionality

All five models of ASC reviewed here capture the assumption of multidimensionality as proposed by Shavelson et al. (1976)—that is, all models differentiate between separate ASC facets representing different domains.

Hierarchy

The assumption of hierarchy as proposed by Shavelson et al. (1976) is included in all of the reviewed models except for the first-order factor model—they all include some form of superordinate constructs. However, the models differ in how they construe the hierarchical nature of ASC. The higher-order factor model includes an HGASC that is located at the apex of the ASC hierarchy and combines the common variance of the first-order ASC factors. The nested Marsh/Shavelson model and the bifactor-ESEM representation include a G-factor that combines the common variance of all manifest ASC indicators (items). The Marsh/Shavelson model includes two higher-order factors (i.e., a higher-order math ASC and a higher-order verbal ASC). As the higher-order factors in the Marsh/Shavelson model are domain-specific, only the higher-order factor model, the nested Marsh/Shavelson model, and the bifactor-ESEM representation meet the original assumption regarding one hierarchically superordinate domain-unspecific, general ASC construct.

Interpretation of Domain-Specific ASCs

All the models considered here reflect a multidimensional nature of ASC and thus include domain-specific ASCs. However, the domain-specific ASCs bear different meanings in the ASC models. In the higher-order factor model, the domain-specific ASC factors capture the residual variance not explained by the HGASC. In the nested Marsh/Shavelson model, the domain-specific ASC factors capture the residual variance not explained by the G-factor. Hence, in these two models, the domain-specific ASCs are interpreted against the background of controlling for a hierarchically superordinate ASC construct (i.e., HGASC or the G-factor). In the Marsh/Shavelson model, the domain-specific ASCs capture the residual variance not explained by the higher-order math and/or the higher-order verbal ASCs. In the bifactor-ESEM representation, the domain-specific ASCs are controlled for when estimating a hierarchically superordinate construct of general ASC (by stating a G-factor) and the other domain-specific ASCs and GASC (by allowing item cross-loadings and stating an S-factor for GASC). Therefore, in all of these

TABLE 1
Theoretical and psychometric characteristics of five central ASC models

Model	Within-network analyses							Between-network analyses: Correlations between ASC and achievement
	Structural characteristics		Higher-order general ASC		Domain-specific ASCs		GASC	
	Multidimensionality	Hierarchy	Interpretation	Invariance of meaning	Interpretation	Conceptual closeness		
HO	Yes	Yes	Common variance across all ASC items (higher-order factor)	No	Residualized for the higher-order factor	No	Residualized for the higher-order factor	Restrictions needed ^a
M/S	Yes	Yes	NA	No	Residualized for higher-order math and verbal factors	Yes: Pattern of factor loadings	Residualized for higher-order math and verbal factors	Restrictions needed ^a
NM/S	Yes	Yes	Common variance of all domain-specific ASC items with GASC (G-factor)	Yes	Residualized for the G-factor	Yes: Pattern of residualized factor correlations	NA	Partial correlation (controlling for the G-factor)
Bi-ESEM	Yes	Yes	Common variance across all ASC items (G-factor)	No	Residualized for the G-factor, GASC, and other domain-specific ASCs	Yes: Pattern of nontarget factor loadings	Residualized for the G-factor and domain-specific ASCs	Partial correlation (controlling for the G-factor, GASC, and other domain-specific ASCs)
FO	Yes	No	NA	NA	First-order factors	Yes: Pattern of factor correlations	First-order factors	Zero-order correlations

Note. ASC = academic self-concept; HO = higher-order factor model; M/S = Marsh/Shavelson model; NM/S = nested Marsh/Shavelson model; Bi-ESEM = bifactor representation using exploratory structural equation modeling; FO = first-order factor model; NA = not applicable; GASC = general academic self-concept.
^aOne model parameter (e.g., factor correlation) needs to be fixed to a predetermined value (e.g., zero).

models, individual students' ratings on domain-specific ASCs provide information about students' profiles (strengths and weaknesses), while keeping constant the level of one (in the higher-order factor model, nested Marsh/Shavelson model, and bifactor-ESEM representation) or two (in the Marsh/Shavelson model) hierarchically superordinate ASC constructs. In the first-order factor model, no hierarchically superordinate construct of ASC is included and cross-loadings are not allowed. Hence, the interpretation of domain-specific ASCs is more complex: Here, the variance of domain-specific ASCs reflects a mixture of the variances of domain-specific ASCs as well as superordinate ASC constructs.

Inspired by the Marsh/Shavelson model (Marsh, 1990b) and the related idea of a math-verbal continuum of domain-specific ASCs (Marsh et al., 1988, 2015), researchers have examined the conceptual closeness of domain-specific ASCs. The models of ASC outlined here differ as to whether and how they depict conceptual closeness among domain-specific ASCs. The Marsh/Shavelson model reflects the conceptual closeness of domain-specific ASCs by the pattern of factor loadings of domain-specific ASCs on the higher-order math and higher-order verbal ASC factors. In the nested Marsh/Shavelson and the first-order factor model, the pattern of correlations among the domain-specific ASCs allows for examining the conceptual closeness of domain-specific ASCs. Importantly, in the nested Marsh/Shavelson model but not in the first-order factor model, these correlations are controlled for the variance attributable to the G-factor. In the nested Marsh/Shavelson model, the correlations among domain-specific ASCs can therefore be clearly attributed to the common variance among domain-specific ASCs. By contrast, in the first-order factor model, the correlations among domain-specific ASCs may partly reflect the common variance attributable to a superordinate ASC in addition to the common variance attributable to domain-specific ASCs. In the bifactor-ESEM representation, the pattern of cross-loadings of items on ASC factors other than the target ASC factors reflects the conceptual overlap between ASC domains. The higher-order factor model cannot provide information on the conceptual closeness among domain-specific ASCs, because all domain-specific ASCs load on the HGASC only.

Interpretation of GASC

The different ASC models conceptualize the directly assessed GASC in different ways. In the higher-order factor model, GASC is a first-order factor that captures the common variance of all items assessing GASC and that is assumed to load on the higher-order factor of HGASC. In the bifactor-ESEM representation, GASC items load on a separate S-factor and on the G-factor (and may additionally have cross-loadings on the other S-factors). In the Marsh/Shavelson model, GASC is a first-order factor that loads on both the higher-order math and verbal ASC factors. GASC is thus residualized for both higher-order factors. In other words, in the higher-order factor model, the bifactor-ESEM representation, and the Marsh/Shavelson model, GASC items explain residual variance of domain-unspecific ASC ratings that is not captured by hierarchically superior constructs (i.e., HGASC, the G-factor, or higher-order math and verbal ASCs) and other domain-specific ASCs (in the bifactor-ESEM representation). By contrast, the nested Marsh/Shavelson model does not include an S-factor for the GASC items,

which are only used to define a superordinate general ASC (the G-factor). Likewise, in the first-order factor model, GASC is a separate first-order factor with correlations to the other domain-specific ASCs.

Invariance of the Meaning of Hierarchically Superordinate ASC Constructs

The different conceptualizations of GASC outlined above affect the interpretation of hierarchically superordinate constructs. In the higher-order factor model, the bifactor-ESEM representation, and the Marsh/Shavelson model, the meaning of the hierarchically superior constructs is linked to the ASC domains included in the model because the corresponding higher-order factors are defined by the common variance across domain-specific ASCs and GASC. In contrast, in the nested Marsh/Shavelson model, as there is no S-factor for GASC, the G-factor primarily depicts the common variance among GASC items and therefore retains its meaning irrespective of the other ASC domains included in the model (Eid et al., 2017).

Correlations Between ASC and Achievement

In the nested Marsh/Shavelson model, one can probe for the correlations between the G-factor and outcome variables and between domain-specific ASCs (S-factors) and outcome variables. The correlations obtained between domain-specific ASCs (S-factors) and outcome variables can be interpreted as semipartial correlations that are controlled for the G-factor. This model does not permit testing the correlations between GASC and outcome variables because there is no S-factor for GASC. The bifactor-ESEM representation enables testing the correlations between outcome variables and the G-factor, the S-factors for domain-specific ASCs, and the S-factor for GASC. The resulting correlations for the S-factors are controlled not only for the G-factor as a hierarchically superordinate construct but also for shared variance with other ASCs given possible item cross-loadings across the S-factors for domain-specific ASCs and GASC (Asphahourov & Muthen, 2009; Morin et al., 2016). The first-order factor model allows researchers to examine the correlations between all domain-specific ASCs as well as GASC and outcome variables. When interpreting the correlations, one should keep in mind that the variances of the ASC factors represent a combination of the variance that is specific to the ASC considered and shared with other ASCs included in the model. Consequently, estimates of correlations with outcome variables obtained from the first-order factor model reflect correlations that are specific to a certain ASC but also due to shared variance across ASCs.

In the higher-order factor model and the Marsh/Shavelson model, one typically considers only the relations of the higher-order factors to outcome variables. When studying correlations between first-order ASCs and outcome variables, the higher-order factor model and the Marsh/Shavelson model suffer from an inherent psychometric restriction, that is, the proportionality constraint (Brunner et al., 2012; Gignac, 2016; see also Chen et al., 2006; Schmiedek & Li, 2004). This constraint affects the proportion of variance in the item scores explained by higher-order and first-order ASC constructs. Specifically, the ratio of variance attributable to the first-order ASCs to variance attributable to the HGASC (in the higher-order factor model) or to the higher-order math and verbal ASCs (in the Marsh/Shavelson model) is constrained to be the same across a given set of ASC items. The estimated relations of first-order and higher-order ASCs (HGASC or

higher-order math and verbal ASCs) and outcome variables are thus linearly dependent (Schmiedek & Li, 2004). In other words, the proportionality constraint limits the value of these higher-order factor models in providing insights into the relation between first-order factors (domain-specific ASCs or GASC) and outcome variables. If the relations between first-order factors and outcome variables are of interest beyond the relations between higher-order factors and outcome variables, one has to rely on highly restrictive model assumptions. Specifically, the size of one model parameter such as a correlation between a first-order factor and the outcome variable has to be fixed to a predetermined value (usually zero) in order to achieve model identification. This model constraint may distort the results when the predetermined value for the ASC-outcome relation does not reflect the true empirical relation (Brunner et al., 2012; Christensen et al., 2001).

Part 2: Empirical Illustration

In Part 1, we provided an overview of five central models of the structure of ASC. These models have been developed and tested using data sets that differ regarding the domain-specific ASCs included and the student samples used (e.g., age and origin of the samples). To our knowledge, no study has yet used the same data set to empirically test the models described above. Our overview of the different ASC models shows that an empirical evaluation of these models has to rely on a data set that includes a measure for GASC as well as measures for a broad variety of domain-specific ASCs. Therefore, we implemented the models using a large data set obtained from secondary school students in Germany with a measure for GASC as well as measures for seven domain-specific ASCs. We scrutinized and compared the properties of these models with within-network analyses, by considering the resulting structural characteristics, and then with between-network analyses, by examining how the obtained correlations between ASCs and academic achievement—operationalized by school grades—vary depending on the structural model of ASC applied.

Method

Sample

On behalf of the 16 federal states in Germany, the Institute for Educational Quality Improvement (IQB) conducts regular, sample-based, large-scale assessment studies to monitor the German educational system and compare student achievement across the federal states (e.g., Pant et al., 2013).³ We used data from a large field trial to pretest achievement test items to be used in these assessment studies. The field trial was conducted in 2014 with a total sample of 3,258 10th-grade students. As is common in large-scale assessments, different test booklets and student questionnaires were randomly assigned to the students. The versions were randomized at the class level (i.e., all students from a class received the same questionnaire). Our analyses are based on the random subsample of $N = 1,232$ students ($N = 587$ [47.6%] boys and $N = 644$ [52.3%] girls, $N = 1$ unreported) who received a questionnaire version with measures of ASC. Students' age ranged from 14 to 19 years with a mean age of 15.43 years ($SD = 0.65$). The students came from 63 classes distributed across 48 schools. All common school types of the German secondary school system were included with a (slight) oversampling

(50.8%) of the academic track (Gymnasium). The majority of students ($N = 1,228$; 87.9%) had German as their native language, $N = 145$ (11.8%) students did not have German as their native language, and respective information was missing for $N = 4$ (0.3%) students. Participation was mandatory at the school level but not at the student level. The participation rate at the student level was 79.6%. Students were not rewarded or graded for participation. Parental consent was given for all students in the present sample.

Measures

Students' ASCs in the domains of German (students' main language of instruction), math, physics, English (a foreign language taught in all secondary schools in the German educational system, most commonly as students' first foreign language), chemistry, biology, and history were measured with four items each. The items were worded identically across the seven domains: "I learn quickly in [domain]"; "I have always been good at [domain]"; "Things in [domain] are easy for me"; "It is easy for me to understand new things in [domain]." The GASC scale consisted of three items that asked the students for their self-perceptions of school competence unrelated to a specific domain: "I learn things quickly in most school subjects"; "I have always been good at school"; "Things in most school subjects are easy for me." Students responded to all items on a 5-point Likert-type scale ranging from *does not apply at all* to *completely applies*. Hence, high ratings consistently represented high levels of ASCs in all domains. The reliability estimates in terms of Cronbach's alpha were good for all scales ($\alpha = .83$ to $.95$; Supplemental Table S1 in the online version of the journal).

Students' school grades related to the same domains as assessed for the domain-specific ASCs—that is, school grades in German, math, physics, English, chemistry, biology, and history (for descriptive statistics, see Supplemental Table S1 in the online version of the journal). The school grades were reported by the school officials based on the latest school report (end of the 9th-grade school report obtained in summer 2014). In Germany, school grades range from 1 to 6, with 1 representing the best grade. For ease of interpretation, school grades were reverse coded for all analyses so that higher numbers represented better achievement.

Statistical Analyses

The analyses were all conducted within the SEM framework using Mplus 8.2 (Muthén & Muthén, 1998–2019). For all models, we used the robust maximum likelihood estimator that may account for nonnormally distributed manifest variables (Hox et al., 2010). As the data had a hierarchical structure (students nested in classes), all models were implemented by using the Mplus option "type=complex," using students' classes as cluster variables. This option corrects for possibly biased standard errors, which can result when not considering the hierarchical structure of the data (Stapleton, 2006). In all models, we further included correlated uniquenesses between parallel-worded ASC items across domains to account for common variance (Marsh et al., 2013). To obtain model identification, the unstandardized loading of the first item on a factor was set to 1 in the models relying on the CFA framework (i.e., the higher-order factor model, the Marsh/Shavelson model, the nested Marsh/Shavelson model, and the

first-order factor model). In the bifactor-ESEM representation, we used target rotation. Here, each item was assumed to have target loadings on the G-factor as well as on the S-factor representing the matching domain-specific ASC (or GASC). All items also had nontarget loadings on all remaining S-factors for domain-specific ASCs or GASC.

In a first step, we analyzed the structural models of ASC (within-network analyses; Byrne, 1996); in a second step, we extended these models by including school grades to examine the correlations between ASC and achievement as an outcome variable (between-network analyses; Byrne, 1996). We used the school grades as single-item indicators to operationalize domain-specific achievement factors. The variance of the residual term of each achievement factor was fixed to zero. We inspected the zero-order correlations between the domain-specific school grades and ASC facets. The Mplus inputs for each ASC model are reported in the supplements in the online version of the journal.

Three students had missing values on all ASC variables and had thus to be excluded from the analyses that focused on the structure of ASC only; these students could, however, be included in the analyses that focused on the relations between ASCs and school grades. On the item level, missingness for the ASC items was very low (range from 0.4% to 1.2%). For the school grades, the amount of missing data was higher and ranged from 2.3% to 20.9%. The major reason for missing values on the school grades was that single subjects were not taught or graded in a specific school, school type, or federal state in Grade 10, which can be considered as a missing at random (MAR) process. Still, the considered school subjects are part of the standard curriculum for secondary schools and thus all students were taught in these school subjects during their school career at some time. Missing values on all variables were handled by the full information maximum likelihood (FIML) approach implemented in Mplus. The FIML approach is known to be reliable and to lead to unbiased parameter estimates when handling data that are MAR or missing completely at random (Enders, 2010; Graham, 2009). The assumption of MAR cannot be empirically tested (Schafer & Graham, 2002), but serious violations of the assumption of MAR are relatively rare (Graham et al., 1997; Schafer & Graham, 2002). School grades from other subjects and domain-specific ASCs were included in the analyses; school grades of different school subjects are highly correlated, and ASCs and school grades of matching domains are also highly correlated (Möller et al., 2009, 2020). Thus, even if the value of a certain school grade drove the missing data process (i.e., data are not missing at random), this process could be approximated by including school grades in other subjects or students' ASCs in different subjects in the FIML estimation process. Simulation studies have shown that including such powerful covariates in the estimation process helps reduce bias in model parameters even when data are not missing at random (Collins et al., 2001). Taken together, missing data may not impose severe threats to the validity of our findings on the comparison of the structural models of ASC or on the pattern of relations between ASCs and achievement.

Two main criteria were applied for model fit evaluation. First, we referred to several commonly accepted descriptive goodness-of-fit indices (Marsh et al., 2004): the comparative fit index (CFI), the root mean square error of approximation

(RMSEA), and the standardized root mean square residual (SRMR). For the CFI, values between .90 and .95 are commonly accepted as indicators of good model fit, although some authors suggest a stricter criterion of .95 (e.g., Hu & Bentler, 1998). Concerning the RMSEA, values below .05 are indicative of a close fit, values between .05 and .08 are indicative of a reasonable fit, and values greater than .10 are indicative of a poor fit (Browne & Cudeck, 1993). For the SRMR, values below .05 indicate good model fit (Diamantopoulos & Siguaw, 2000), while cutoff values of .08 (Hu & Bentler, 1998) and even of .10 (Kline, 2005) are still accepted as adequate. Second, we inspected the resulting model parameter estimates and how well they aligned with theoretical assumptions of ASC (Bollen, 1989; Kline, 2005; Marsh et al., 2004; West et al., 2012). Particularly, we evaluated the statistical significance and size of the standardized factor loadings (λ). Following Floyd and Widaman (1995), we considered standardized factor loadings of $\lambda \geq .30$ to be substantial. For cross-loadings resulting from the bifactor-ESEM representation, we followed Mai et al. (2018), who proposed that cross-loadings of $\lambda \geq .10$ are nonignorable. In addition, we evaluated factor correlations among the ASCs and the correlations between ASCs and academic achievement operationalized by school grades to judge the theoretical and empirical adequacy of the ASC models.

Results

For each ASC model, we first report the key results obtained for the structural model itself. Second, we report the findings regarding the correlations between ASCs and school grades.

Higher-Order Factor Model

Most goodness-of-fit indices obtained for the higher-order factor model with and without school grades were within an acceptable range, except for the SRMR that was greater than .10 (Table 2). All domain-specific first-order ASCs and GASC were well-defined as is evident from substantial and statistically significant positive factor loadings of the items on their matching first-order factors (Table 3). Among the first-order factors, GASC displayed the highest loading on the HGASC ($\lambda = .79$). The domain-specific first-order factors for English ($\lambda = .19$), German ($\lambda = .21$), and history ASCs ($\lambda = .35$) were relatively weakly related to the HGASC, whereas biology ($\lambda = .54$), math ($\lambda = .64$), chemistry ($\lambda = .72$), and physics ($\lambda = .72$) ASCs demonstrated substantial loadings on the HGASC.

The HGASC was positively related to academic achievement as measured by school grades (Table 4; see also Supplemental Table S2 in the online version of the journal). The correlations ranged from $r = .20$ for the English grade to $r = .55$ for the chemistry grade. The pattern of correlations showed that the HGASC was more strongly related to grades obtained in math/science subjects (math: $r = .47$; physics: $r = .50$; chemistry: $r = .55$) than to school grades in verbal subjects (German: $r = .26$; English: $r = .20$).

Marsh/Shavelson Model

The Marsh/Shavelson model fitted the data well both with and without school grades (Table 2). The first-order, domain-specific ASC factors and GASC were

TABLE 2
Goodness-of-fit indices

Model description	χ^2	<i>df</i>	CFI	RMSEA	SRMR
Models without school grades					
Higher-order factor model	1411.393	321	.960	.053	.104
Marsh/Shavelson model	791.001	317	.982	.035	.044
Nested Marsh/Shavelson model	552.248	280	.990	.028	.020
Bifactor-ESEM representation	223.606	117	.996	.027	.005
First-order factor model	618.927	301	.988	.029	.025
Models with school grades					
Higher-order factor model	3288.081	531	.916	.065	.107
Marsh/Shavelson model	2198.234	520	.949	.051	.058
Nested Marsh/Shavelson model	986.127	441	.983	.032	.024
Bifactor-ESEM representation	594.596	271	.990	.031	.008
First-order factor model	1062.012	462	.982	.032	.028

Note. All χ^2 were statistically significant ($p < .05$). All models were conducted with the robust maximum likelihood estimator. ESEM = exploratory structural equation modeling; CFI = comparative fit index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual.

well-defined given the statistically significant and substantial positive loadings of the items on the corresponding factors (Table 3). Both higher-order factors were also well-defined: The higher-order math ASC was defined by math-like ASCs such as math ASC ($\lambda = .73$), chemistry ASC ($\lambda = .74$), and physics ASC ($\lambda = .76$); the higher-order verbal ASC was defined by verbal-like ASCs such as German ASC ($\lambda = .78$) and English ASC ($\lambda = .49$). The first-order GASC factor showed substantial positive loadings of similar size on both higher-order factors (higher-order math ASC: $\lambda = .68$; higher-order verbal ASC: $\lambda = .64$). History ASC and biology ASC also displayed substantial positive loadings on both higher-order factors. History ASC showed a higher loading on the higher-order verbal ASC ($\lambda = .44$) than on the higher-order math ASC ($\lambda = .23$). Biology ASC showed a similarly high loading on the higher-order verbal ASC ($\lambda = .34$) and higher-order math ASC ($\lambda = .45$). The correlation between the higher-order math ASC and the higher-order verbal ASC factors was not statistically significant ($r = -.05$).

The higher-order math ASC was more strongly related to school grades in math/science subjects such as math ($r = .50$), physics ($r = .49$), and chemistry ($r = .54$) than to school grades in verbal subjects (German: $r = .14$; English: $r = .08$; Table 4; see also Supplemental Table S3 in the online version of the journal). The higher-order verbal ASC, in turn, was more strongly related to school grades in verbal subjects such as German ($r = .47$) and English ($r = .47$) than to grades obtained in math/science subjects (math: $r = .05$; physics: $r = .12$; chemistry: $r = .05$). Both higher-order ASC factors were positively related to the biology grade ($r = .36$ for the higher-order math ASC, and $r = .21$ for the

TABLE 3
Standardized factor loadings (p values in parentheses) of the ASC models

Domain	Higher-order factor model		Marsh/Shavelson model			Nested Marsh/Shavelson model		First-order factor model
	First-order loadings	Second-order loadings	First-order loadings	Second-order loadings on math ASC	Second-order loadings on verbal ASC	First-order loadings	Loadings on the G-factor	First-order loadings
Math ASC		.636 (.000)/ .671 (.000)		.727 (.000)/ .746 (.000)				
Item 1	.891 (.000)/ .893 (.000)		.890 (.000)/ .892 (.000)			.776 (.000)/ .775 (.000)	.442 (.000)/ .448 (.000)	.891 (.000)/ .893 (.000)
Item 2	.803 (.000)/ .806 (.000)		.804 (.000)/ .807 (.000)			.678 (.000)/ .680 (.000)	.439 (.000)/ .446 (.000)	.805 (.000)/ .810 (.000)
Item 3	.921 (.000)/ .921 (.000)		.921(.000)/ .920 (.000)			.796 (.000)/ .798 (.000)	.463 (.000)/ .457 (.000)	.921 (.000)/ .920 (.000)
Item 4	.919 (.000)/ .918 (.000)		.919 (.000)/ .917 (.000)			.767 (.000)/ .767 (.000)	.506 (.000)/ .499 (.000)	.918 (.000)/ .915 (.000)
German ASC		.207 (.020)/ .144 (.096)			.781 (.000)/ .752 (.000)			
Item 1	.843 (.000)/ .841 (.000)		.853 (.000)/ .852 (.000)			.748 (.000)/ .748 (.000)	.410 (.000)/ .410 (.000)	.853 (.000)/ .852 (.000)
Item 2	.779 (.000)/ .779 (.000)		.792 (.000)/ .798 (.000)			.716 (.000)/ .719 (.000)	.338 (.000)/ .348 (.000)	.792 (.000)/ .798 (.000)
Item 3	.861 (.000)/ .859 (.000)		.861 (.000)/ .857 (.000)			.768 (.000)/ .767 (.000)	.384 (.000)/ .376 (.000)	.860 (.000)/ .856 (.000)
Item 4	.849 (.000)/ .848 (.000)		.855 (.000)/ .855 (.000)			.759 (.000)/ .757 (.000)	.399 (.000)/ .403 (.000)	.857 (.000)/ .857 (.000)
English ASC		.187 (.006)/ .133 (.053)			.492 (.000)/ .545 (.000)			
Item 1	.921 (.000)/ .921 (.000)		.922 (.000)/ .922 (.000)			.852 (.000)/ .852 (.000)	.354 (.000)/ .357 (.000)	.922 (.000)/ .923 (.000)
Item 2	.854 (.000)/ .853 (.000)		.860 (.000)/ .861 (.000)			.804 (.000)/ .807 (.000)	.305 (.000)/ .307 (.000)	.860 (.000)/ .863 (.000)
Item 3	.922 (.000)/ .921 (.000)		.921 (.000)/ .920 (.000)			.850 (.000)/ .849 (.000)	.353 (.000)/ .351 (.000)	.921 (.000)/ .919 (.000)
Item 4	.910 (.000)/ .910 (.000)		.912 (.000)/ .913 (.000)			.833 (.000)/ .831 (.000)	.371 (.000)/ .374 (.000)	.912 (.000)/ .911 (.000)
Physics ASC		.721 (.000)/ .725 (.000)		.764 (.000)/ .719 (.000)				
Item 1	.922 (.000)/ .922 (.000)		.922 (.000)/ .922 (.000)			.801 (.000)/ .801 (.000)	.455 (.000)/ .453 (.000)	.922 (.000)/ .921 (.000)
Item 2	.873 (.000)/ .874 (.000)		.872 (.000)/ .873 (.000)			.731 (.000)/ .731 (.000)	.479 (.000)/ .486 (.000)	.871 (.000)/ .874 (.000)
Item 3	.913 (.000)/ .911 (.000)		.912 (.000)/ .911 (.000)			.821 (.000)/ .823 (.000)	.405 (.000)/ .400 (.000)	.913 (.000)/ .912 (.000)
Item 4	.902 (.000)/ .901 (.000)		.901 (.000)/ .900 (.000)			.802 (.000)/ .803 (.000)	.418 (.000)/ .414 (.000)	.902 (.000)/ .901 (.000)
Biology ASC		.544 (.000)/ .519 (.000)		.449 (.000)/ .449 (.000)	.341 (.000)/ .305 (.000)			
Item 1	.878 (.000)/ .878 (.000)		.881 (.000)/ .881 (.000)			.776 (.000)/ .775 (.000)	.416 (.000)/ .418 (.000)	.881 (.000)/ .880 (.000)
Item 2	.849 (.000)/ .850 (.000)		.851 (.000)/ .852 (.000)			.761 (.000)/ .760 (.000)	.389 (.000)/ .396 (.000)	.852 (.000)/ .855 (.000)
Item 3	.874 (.000)/ .872 (.000)		.873 (.000)/ .871 (.000)			.775 (.000)/ .778 (.000)	.400 (.000)/ .392 (.000)	.873 (.000)/ .872 (.000)
Item 4	.877 (.000)/ .876 (.000)		.878 (.000)/ .877 (.000)			.752 (.000)/ .752 (.000)	.457 (.000)/ .456 (.000)	.878 (.000)/ .878 (.000)
History ASC		.354 (.000)/ .309 (.000)		.230 (.000)/ .209 (.000)	.435 (.000)/ .426 (.000)			
Item 1	.892 (.000)/ .891 (.000)		.894 (.000)/ .894 (.000)			.830 (.000)/ .831 (.000)	.333 (.000)/ .332 (.000)	.894 (.000)/ .895 (.000)
Item 2	.858 (.000)/ .858 (.000)		.858 (.000)/ .859 (.000)			.799 (.000)/ .800 (.000)	.315 (.000)/ .319 (.000)	.859 (.000)/ .861 (.000)
Item 3	.908 (.000)/ .907 (.000)		.909 (.000)/ .907 (.000)			.849 (.000)/ .852 (.000)	.327 (.000)/ .317 (.000)	.910 (.000)/ .909 (.000)
Item 4	.888 (.000)/ .887 (.000)		.889 (.000)/ .889 (.000)			.816 (.000)/ .814 (.000)	.357 (.000)/ .357 (.000)	.890 (.000)/ .888 (.000)

(continued)

TABLE 3 (continued)

Domain	Higher-order factor model		Marsh/Shavelson model			Nested Marsh/Shavelson model		First-order factor model
	First-order loadings	Second-order loadings	First-order loadings	Second-order loadings on math ASC	Second-order loadings on verbal ASC	First-order loadings	Loadings on the G-factor	First-order loadings
Chemistry ASC		.717 (.000)/ .736 (.000)		.743 (.000)/ .728 (.000)				
Item 1	.920 (.000)/ .920 (.000)		.919 (.000)/ .919 (.000)			.793 (.000)/ .795 (.000)	.465 (.000)/ .459 (.000)	.920 (.000)/ .919 (.000)
Item 2	.897 (.000)/ .898 (.000)		.896 (.000)/ .898 (.000)			.763 (.000)/ .764 (.000)	.472 (.000)/ .476 (.000)	.897 (.000)/ .899 (.000)
Item 3	.925 (.000)/ .924 (.000)		.924 (.000)/ .923 (.000)			.825 (.000)/ .827 (.000)	.422 (.000)/ .415 (.000)	.924 (.000)/ .923 (.000)
Item 4	.918 (.000)/ .918 (.000)		.918 (.000)/ .917 (.000)			.796 (.000)/ .799 (.000)	.458 (.000)/ .452 (.000)	.918 (.000)/ .918 (.000)
General ASC		.786 (.000)/ .769 (.000)		.684 (.000)/ .719 (.000)	.638 (.000)/ .660 (.000)			
Item 1	.833 (.000)/ .825 (.000)		.853 (.000)/ .847 (.000)			.849 (.000)/ .839 (.000)	.848 (.000)/ .840 (.000)	.849 (.000)/ .839 (.000)
Item 2	.699 (.000)/ .709 (.000)		.706 (.000)/ .726 (.000)			.706 (.000)/ .736 (.000)	.708 (.000)/ .735 (.000)	.706 (.000)/ .736 (.000)
Item 3	.819 (.000)/ .817 (.000)		.807 (.000)/ .803 (.000)			.804 (.000)/ .791 (.000)	.812 (.000)/ .802 (.000)	.804 (.000)/ .791 (.000)

Note. The parameter estimates before the slash were obtained from the models without school grades; the parameter estimates after the slash were obtained from the models including school grades. ASC = academic self-concept.

higher-order verbal ASC) and the history grade ($r = .24$ for the higher-order math ASC, and $r = .35$ for the higher-order verbal ASC).

Nested Marsh/Shavelson Model

The nested Marsh/Shavelson model provided a good fit to the data both with and without school grades (Table 2). The loadings of the domain-specific ASC items on their corresponding S-factors for domain-specific ASCs were all positive and of substantial size (range from $\lambda = .68$ for math ASC to $\lambda = .85$ for English ASC; Table 3). The G-factor was well-defined as is evident from the substantial positive loadings of the GASC items on the G-factor ($\lambda_s = .71$ to $\lambda = .85$). The items measuring domain-specific ASCs also showed positive loadings on the G-factor of a similar size across the different ASC domains ($\lambda_s = .31$ to $\lambda = .51$). The factor correlations (Table 5) showed a clear separation between math-like and verbal-like ASCs. For instance, the correlation between math and German ASCs was $r = -.47$, the correlation between math and English ASCs was $r = -.31$, and the correlation between German and physics ASCs was $r = -.32$. High positive correlations were observed between math-like ASCs such as math and physics ASCs ($r = .40$), math and chemistry ASCs ($r = .31$), or physics and chemistry ASCs ($r = .47$). The two language ASCs (i.e., German and English) were also positively correlated ($r = .20$).

Each domain-specific ASC demonstrated the highest correlations with the school grade of the matching domain (e.g., math ASC and math grade: $r = .33$; Table 4; see also Supplemental Table S4 in the online version of the journal). Moreover, correlations between math-like ASCs and school grades in math/science subjects were positive and statistically significant (e.g., math ASC and

TABLE 4
Correlations with school grades (p values in parentheses)

Model	MGR	GGR	PGR	EGR	CGR	HGR	BGR
Higher-order factor model							
HGASC	.467 (.000)	.263 (.000)	.500 (.000)	.203 (.000)	.548 (.000)	.340 (.000)	.407 (.000)
Marsh/Shavelson model							
Math	.500 (.000)	.142 (.000)	.489 (.000)	.079 (.025)	.544 (.000)	.241 (.000)	.358 (.000)
Verbal	.045 (.236)	.470 (.000)	.120 (.004)	.471 (.000)	.053 (.160)	.353 (.000)	.207 (.000)
Nested Marsh/Shavelson model							
MSC	.326 (.000)	-.120 (.000)	.121 (.001)	-.156 (.000)	.145 (.000)	-.075 (.008)	-.015 (.642)
GSC	-.227 (.000)	.250 (.000)	-.176 (.000)	.095 (.002)	-.173 (.000)	.037 (.236)	-.088 (.015)
PSC	.046 (.196)	-.150 (.000)	.207 (.000)	-.167 (.000)	.119 (.001)	-.044 (.202)	.015 (.641)
ESC	-.146 (.000)	-.034 (.231)	-.139 (.000)	.361 (.000)	-.189 (.000)	-.047 (.080)	-.115 (.000)
CSC	.066 (.033)	-.117 (.001)	.050 (.177)	-.137 (.000)	.326 (.000)	-.036 (.245)	.066 (.040)
HSC	-.188 (.000)	-.030 (.316)	-.088 (.002)	-.075 (.008)	-.123 (.000)	.252 (.000)	-.047 (.079)
BSC	-.115 (.001)	-.114 (.001)	-.030 (.308)	-.145 (.000)	.021 (.554)	.029 (.412)	.199 (.000)
G-factor	.449 (.000)	.442 (.000)	.479 (.000)	.396 (.000)	.449 (.000)	.418 (.000)	.427 (.000)

(continued)

TABLE 4 (continued)

Model	MGR	GGR	PGR	EGR	CGR	HGR	BGR
Bifactor-ESEM representation							
MSC	.380 (.000)	.063 (.267)	.148 (.013)	.060 (.224)	.109 (.086)	.020 (.671)	.027 (.653)
GSC	-.025 (.475)	.378 (.000)	-.002 (.961)	.160 (.000)	-.003 (.930)	.092 (.006)	.020 (.631)
PSC	.007 (.913)	-.024 (.676)	.235 (.000)	-.014 (.791)	.019 (.755)	-.028 (.587)	.005 (.930)
ESC	-.005 (.887)	.037 (.265)	-.010 (.746)	.441 (.000)	-.068 (.023)	.026 (.321)	-.017 (.625)
CSC	.054 (.450)	.019 (.776)	.023 (.743)	.046 (.430)	.304 (.000)	.002 (.963)	.054 (.358)
HSC	-.104 (.003)	.013 (.738)	-.029 (.457)	-.007 (.849)	-.083 (.030)	.290 (.000)	-.021 (.617)
BSC	-.045 (.379)	-.037 (.443)	.024 (.552)	-.033 (.447)	.014 (.767)	.049 (.229)	.240 (.000)
GASC	.363 (.000)	.397 (.000)	.430 (.000)	.382 (.000)	.326 (.000)	.322 (.000)	.348 (.000)
G-factor	.334 (.000)	.211 (.001)	.361 (.000)	.138 (.019)	.413 (.000)	.290 (.000)	.328 (.000)
First-order factor model							
MSC	.511 (.000)	.125 (.000)	.350 (.000)	.070 (.032)	.357 (.000)	.151 (.000)	.207 (.000)
GSC	.006 (.863)	.426 (.000)	.064 (.073)	.267 (.000)	.053 (.123)	.226 (.000)	.119 (.001)
PSC	.255 (.000)	.079 (.018)	.410 (.000)	.042 (.152)	.319 (.000)	.162 (.000)	.217 (.000)
ESC	.040 (.165)	.141 (.000)	.058 (.050)	.486 (.000)	.000 (.989)	.119 (.000)	.060 (.060)
CSC	.277 (.000)	.114 (.001)	.277 (.000)	.074 (.017)	.505 (.000)	.172 (.000)	.267 (.000)
HSC	-.008 (.809)	.137 (.000)	.096 (.007)	.077 (.018)	.052 (.050)	.389 (.000)	.115 (.000)
BSC	.113 (.000)	.111 (.001)	.202 (.000)	.061 (.098)	.233 (.000)	.225 (.000)	.379 (.000)
GASC	.444 (.000)	.438 (.000)	.469 (.000)	.394 (.000)	.444 (.000)	.414 (.000)	.423 (.000)

Note. For the full correlations including the correlations among school grades and academic self-concepts, see the supplements (Supplemental Tables S2 to S6 in the online version of the journal); MGR = math grade; GGR = German grade; PGR = physics grade; EGR = English grade; CGR = chemistry grade; HGR = history grade; BGR = biology grade; HGASC = higher-order factor for general academic self-concept; MATH = higher-order math self-concept; VERBAL = higher-order verbal self-concept; MSC = math self-concept; GSC = German self-concept; PSC = physics self-concept; ESC = English self-concept; CSC = chemistry self-concept; HSC = history self-concept; BSC = biology self-concept; GASC = general academic self-concept; ESEM = exploratory structural equation modeling.

TABLE 5
Factor correlations (with p values in parentheses) of the nested Marsh/Shavelson model and first-order factor model

ASC domain	Math ASC		German ASC		English ASC		Physics ASC		Biology ASC		Chemistry ASC		History ASC	
	NM/S	FO	NM/S	FO	NM/S	FO	NM/S	FO	NM/S	FO	NM/S	FO	NM/S	FO
German ASC	-.470 (.000)	-.119 (.001)												
English ASC	-.311 (.000)	-.045 (.211)	.204 (.000)	.345 (.000)										
Physics ASC	.397 (.000)	.548 (.000)	-.320 (.000)	-.030 (.464)	-.170 (.000)	.047 (.188)								
Biology ASC	-.049 (.147)	.211 (.000)	.058 (.146)	.265 (.000)	-.070 (.099)	.127 (.001)	.151 (.000)	.345 (.000)						
Chemistry ASC	.314 (.000)	.490 (.000)	-.288 (.000)	.004 (.909)	-.217 (.000)	.016 (.666)	.469 (.000)	.595 (.000)	.238 (.000)	.417 (.000)				
History ASC	-.154 (.000)	.073 (.039)	.242 (.000)	.372 (.000)	.027 (.511)	.167 (.000)	.075 (.036)	.240 (.000)	.177 (.000)	.322 (.000)	-.019 (.558)	.168 (.000)		
General ASC	—	.517 (.000)	—	.457 (.000)	—	.387 (.000)	—	.475 (.000)	—	.477 (.000)	—	.493 (.000)	.375 (.000)	
G-factor	.000	—	.000	—	.000	—	.000	—	.000	—	.000	—	.000	—

Note. A G-factor is only modeled in the nested Marsh/Shavelson model; a first-order factor for general ASC is only included in the first-order factor model. NM/S = nested Marsh/Shavelson model; FO = first-order factor model; ASC = academic self-concept.

TABLE 6
Bifactor-ESEM representation: Standardized factor loadings (p values in parentheses)

	Math ASC	German ASC	English ASC	Physics ASC	Biology ASC	History ASC	Chemistry ASC	General ASC	G-factor
Math ASC									
Item 1	.728 (.026)/ .730 (.000)	-.115 (.533)/ -.120 (.000)	-.074 (.312)/ -.075 (.000)	.081 (.942)/ .063 (.002)	-.053 (.881)/ -.057 (.003)	-.078 (.321)/ -.076 (.000)	.031 (.834)/ .034 (.219)	.061 (.712)/ .041 (.024)	.489 (.465)/ .496 (.000)
Item 2	.598 (.237)/ .606 (.000)	-.116 (.627)/ -.121 (.000)	-.071 (.491)/ -.074 (.002)	.076 (.955)/ .064 (.058)	-.037 (.940)/ -.038 (.206)	-.063 (.033)/ -.056 (.030)	.031 (.921)/ .048 (.170)	.093 (.818)/ .136 (.000)	.497 (.541)/ .501 (.000)
Item 3	.696 (.066)/ .681 (.000)	-.142 (.175)/ -.148 (.000)	-.089 (.194)/ -.095 (.000)	.043 (.954)/ .038 (.287)	-.067 (.812)/ -.075 (.001)	-.072 (.512)/ -.079 (.000)	.017 (.952)/ .003 (.927)	.010 (.980)/ -.023 (.411)	.573 (.116)/ .581 (.000)
Item 4	.693 (.183)/ .663 (.000)	-.075 (.757)/ -.089 (.000)	-.068 (.497)/ -.076 (.000)	.109 (.937)/ .059 (.028)	-.053 (.915)/ -.081 (.000)	-.076 (.078)/ -.082 (.000)	.025 (.933)/ .001 (.974)	.023 (.957)/ -.045 (.037)	.576 (.486)/ .612 (.000)
German ASC									
Item 1	-.071 (.482)/ -.075 (.002)	.793 (.000)/ .791 (.000)	.115 (.000)/ .115 (.000)	-.077 (.842)/ -.089 (.000)	.043 (.797)/ .037 (.084)	.145 (.100)/ .141 (.000)	-.058 (.137)/ -.064 (.004)	.129 (.410)/ .089 (.002)	.181 (.172)/ .196 (.000)
Item 2	-.118 (.237)/ -.113 (.000)	.749 (.000)/ .753 (.000)	.120 (.000)/ .122 (.000)	-.093 (.718)/ -.091 (.002)	.028 (.852)/ .028 (.307)	.083 (.103)/ .089 (.000)	-.079 (.168)/ -.065 (.054)	.118 (.541)/ .157 (.000)	.120 (.139)/ .127 (.021)
Item 3	-.069 (.724)/ -.094 (.002)	.819 (.000)/ .803 (.000)	.161 (.000)/ .155 (.000)	-.024 (.978)/ -.072 (.028)	.108 (.736)/ .087 (.000)	.120 (.587)/ .119 (.000)	-.021 (.622)/ -.039 (.211)	.011 (.956)/ -.030 (.247)	.159 (.742)/ .200 (.001)
Item 4	-.131 (.034)/ -.128 (.000)	.784 (.000)/ .784 (.000)	.154 (.000)/ .155 (.000)	-.052 (.857)/ -.053 (.023)	.084 (.523)/ .082 (.000)	.158 (.066)/ .160 (.000)	-.073 (.022)/ -.073 (.007)	.068 (.680)/ .055 (.051)	.187 (.009)/ .194 (.000)
English ASC									
Item 1	-.036 (.514)/ -.039 (.063)	.112 (.102)/ .109 (.000)	.897 (.000)/ .897 (.000)	.015 (.961)/ .004 (.799)	-.007 (.952)/ -.011 (.462)	.055 (.422)/ .055 (.003)	-.041 (.067)/ -.044 (.004)	.074 (.375)/ .053 (.000)	.154 (.199)/ .166 (.000)
Item 2	-.086 (.453)/ -.092 (.001)	.116 (.147)/ .115 (.000)	.836 (.000)/ .837 (.000)	-.041 (.906)/ -.056 (.038)	.012 (.938)/ .005 (.838)	.016 (.663)/ .015 (.515)	-.032 (.694)/ -.035 (.184)	.126 (.377)/ .121 (.001)	.096 (.531)/ .113 (.019)
Item 3	-.070 (.701)/ -.063 (.005)	.105 (.560)/ .110 (.000)	.891 (.000)/ .890 (.000)	-.070 (.927)/ -.030 (.227)	-.005 (.985)/ .007 (.728)	.035 (.822)/ .027 (.182)	-.042 (.563)/ -.041 (.041)	.004 (.947)/ .005 (.827)	.208 (.763)/ .192 (.000)
Item 4	-.044 (.736)/ -.053 (.008)	.162 (.069)/ .156 (.000)	.878 (.000)/ .874 (.000)	-.009 (.982)/ -.034 (.074)	.027 (.877)/ .015 (.499)	.033 (.488)/ .033 (.109)	-.048 (.418)/ -.056 (.015)	.048 (.813)/ .031 (.203)	.170 (.307)/ .191 (.000)
Physics ASC									
Item 1	.035 (.736)/ .085 (.009)	-.105 (.005)/ -.093 (.000)	-.069 (.043)/ -.059 (.005)	.551 (.128)/ .616 (.000)	-.082 (.218)/ -.045 (.086)	-.025 (.749)/ -.002 (.940)	-.035 (.773)/ .029 (.439)	-.100 (.674)/ -.050 (.070)	.728 (.009)/ .667 (.000)
Item 2	.014 (.721)/ .064 (.024)	-.084 (.168)/ -.071 (.000)	-.035 (.498)/ -.027 (.228)	.517 (.063)/ .599 (.000)	-.036 (.334)/ .006 (.775)	-.019 (.874)/ .003 (.910)	-.029 (.891)/ .036 (.286)	-.017 (.827)/ .088 (.052)	.693 (.001)/ .636 (.000)
Item 3	.181 (.918)/ .065 (.142)	-.019 (.991)/ -.010 (.000)	-.011 (.979)/ -.035 (.070)	1.038 (.892)/ .677 (.000)	.125 (.960)/ -.011 (.740)	.021 (.989)/ .021 (.300)	.147 (.664)/ .089 (.029)	.080 (.912)/ -.064 (.084)	.379 (.946)/ .597 (.000)
Item 4	-.009 (.950)/ .023 (.405)	-.124 (.181)/ -.112 (.000)	-.040 (.542)/ -.037 (.060)	.563 (.000)/ .626 (.000)	-.042 (.760)/ -.013 (.551)	-.009 (.946)/ -.002 (.913)	.033 (.912)/ .070 (.013)	-.099 (.289)/ -.075 (.010)	.681 (.000)/ .637 (.000)
Biology ASC									
Item 1	-.046 (.873)/ -.037 (.099)	.075 (.520)/ .074 (.000)	-.003 (.937)/ .000 (.996)	-.012 (.989)/ -.017 (.409)	.738 (.016)/ .739 (.000)	.059 (.602)/ .066 (.000)	.028 (.811)/ .043 (.038)	.022 (.919)/ .019 (.280)	.465 (.306)/ .461 (.000)
Item 2	-.078 (.786)/ -.070 (.005)	.077 (.490)/ .076 (.000)	-.015 (.668)/ -.015 (.459)	-.017 (.984)/ -.014 (.579)	.722 (.017)/ .728 (.000)	.064 (.525)/ .074 (.001)	.015 (.906)/ .035 (.151)	.036 (.877)/ .088 (.000)	.436 (.328)/ .434 (.000)
Item 3	-.027 (.963)/ -.071 (.004)	.050 (.890)/ .029 (.214)	.035 (.740)/ .023 (.242)	.061 (.975)/ -.027 (.387)	.763 (.258)/ .722 (.000)	.041 (.889)/ .038 (.115)	.041 (.860)/ .004 (.869)	.005 (.990)/ -.065 (.011)	.425 (.728)/ .490 (.000)
Item 4	-.046 (.874)/ -.055 (.000)	.083 (.447)/ .075 (.000)	.014 (.713)/ .010 (.626)	.025 (.977)/ .004 (.821)	.716 (.015)/ .705 (.000)	.063 (.557)/ .062 (.001)	.023 (.844)/ .016 (.458)	.015 (.946)/ -.011 (.563)	.500 (.278)/ .514 (.000)
History ASC									
Item 1	-.003 (.994)/ -.037 (.248)	.138 (.672)/ .121 (.000)	.039 (.739)/ .032 (.172)	.089 (.950)/ .017 (.511)	.115 (.835)/ .083 (.000)	.842 (.000)/ .835 (.000)	.025 (.932)/ -.006 (.852)	.106 (.678)/ .042 (.117)	.246 (.756)/ .293 (.000)
Item 2	-.041 (.943)/ -.070 (.071)	.123 (.725)/ .107 (.000)	.032 (.817)/ .025 (.310)	.089 (.953)/ .019 (.641)	.096 (.879)/ .067 (.019)	.792 (.000)/ .797 (.000)	.024 (.953)/ .007 (.885)	.094 (.825)/ .088 (.156)	.253 (.769)/ .297 (.000)
Item 3	-.180 (.896)/ -.096 (.039)	.065 (.962)/ .120 (.000)	.025 (.945)/ .037 (.107)	-.336 (.955)/ -.048 (.420)	-.083 (.966)/ .023 (.534)	.862 (.494)/ .821 (.000)	-.136 (.698)/ -.097 (.042)	-.168 (.707)/ -.090 (.065)	.538 (.907)/ .367 (.000)
Item 4	-.008 (.989)/ -.040 (.071)	.161 (.634)/ .143 (.000)	.054 (.667)/ .047 (.013)	.114 (.939)/ .029 (.224)	.080 (.898)/ .041 (.043)	.808 (.000)/ .802 (.000)	-.019 (.957)/ -.046 (.069)	.048 (.916)/ -.007 (.806)	.288 (.728)/ .341 (.000)
Chemistry ASC									
Item 1	.081 (.877)/ .056 (.008)	-.045 (.887)/ -.064 (.000)	-.029 (.794)/ -.036 (.033)	.156 (.926)/ .094 (.000)	.058 (.920)/ .032 (.106)	-.045 (.735)/ -.049 (.010)	.677 (.012)/ .654 (.000)	.034 (.902)/ -.019 (.323)	.596 (.500)/ .628 (.000)
Item 2	.053 (.932)/ .034 (.063)	-.026 (.941)/ -.040 (.036)	-.062 (.622)/ -.070 (.000)	.134 (.940)/ .078 (.000)	.069 (.914)/ .048 (.001)	-.033 (.737)/ -.029 (.072)	.657 (.074)/ .654 (.000)	.090 (.799)/ .110 (.009)	.589 (.543)/ .620 (.000)
Item 3	-.082 (.934)/ .000 (.993)	-.161 (.886)/ -.109 (.000)	-.094 (.745)/ -.082 (.000)	-.248 (.957)/ .013 (.641)	-.089 (.954)/ .007 (.705)	-.012 (.991)/ -.044 (.009)	.583 (.006)/ .627 (.000)	-.174 (.541)/ -.105 (.000)	.834 (.823)/ .667 (.000)
Item 4	.044 (.952)/ -.004 (.861)	-.059 (.876)/ -.086 (.000)	-.044 (.756)/ -.056 (.003)	.135 (.945)/ .036 (.145)	.072 (.924)/ .023 (.221)	-.053 (.531)/ -.061 (.001)	.648 (.142)/ .609 (.000)	.008 (.989)/ -.080 (.002)	.614 (.556)/ .674 (.000)

(continued)

TABLE 6 (continued)

	Math ASC	German ASC	English ASC	Physics ASC	Biology ASC	History ASC	Chemistry ASC	General ASC	G-factor
General ASC									
Item 1	.133 (.875)/ .064 (.083)	.224 (.520)/ .199 (.000)	.159 (.220)/ .143 (.000)	.045 (.983)/ -.058 (.093)	.083 (.922)/ .020 (.534)	.108 (.001)/ .076 (.003)	.009 (.988)/ -.067 (.052)	.473 (.548)/ .309 (.000)	.613 (.615)/ .702 (.000)
Item 2	.063 (.929)/ .063 (.010)	.182 (.506)/ .178 (.000)	.173 (.087)/ .168 (.000)	.008 (.996)/ -.015 (.591)	.059 (.934)/ .051 (.037)	.001 (.953)/ .012 (.615)	-.049 (.929)/ -.034 (.248)	.509 (.454)/ .628 (.000)	.481 (.632)/ .515 (.000)
Item 3	.071 (.690)/ .028 (.468)	.139 (.652)/ .137 (.000)	.136 (.015)/ .124 (.000)	-.155 (.849)/ -.162 (.000)	-.033 (.802)/ -.059 (.046)	.013 (.980)/ -.026 (.368)	.017 (.968)/ -.054 (.172)	.399 (.464)/ .249 (.000)	.704 (.421)/ .744 (.000)

Note. Target factor loadings are indicated in bold. The parameter estimates before the slash were obtained from the model without school grades; the parameter estimates after the slash were obtained from the model including school grades. ESEM = exploratory structural equation modeling; ASC = academic self-concept.

physics grade: $r = .12$; math ASC and chemistry grade: $r = .15$). Correlations between verbal-like ASCs and school grades in verbal subjects were positive and statistically significant (e.g., German ASC and English grade: $r = .10$) or not statistically significant (English ASC and German grade: $r = -.03$). Negative correlations were observed between math-like ASCs and school grades in verbal subjects (e.g., math ASC and German grade: $r = -.12$; physics ASC and English grade: $r = -.17$) as well as between verbal-like ASCs and school grades in math/science subjects (e.g., German ASC and math grade: $r = -.23$; English ASC and physics grade: $r = -.14$). Finally, the G-factor demonstrated substantial positive relations to school grades in all subjects ($rs = .40$ to $.48$).

Bifactor-ESEM Representation

Despite the very good model fit obtained for the bifactor-ESEM representation (Table 2), an inspection of the factor loadings indicated several estimation problems. Out of 31 target loadings on the S-factors, 11 loadings did not reach statistical significance, although they were substantial in size (i.e., $\lambda \geq .30$; Table 6). For the S-factor of physics ASC, there was one standardized factor loading above 1 that was not statistically significant. The target loadings of the domain-specific ASC items and GASC items on the G-factor were all nonsignificant. Nevertheless, the inclusion of item cross-loadings seemed to be warranted as 52 out of the 217 item cross-loadings were larger than $\lambda > .10$ and therefore nonignorable, although most of these cross-loadings were not statistically significant. Finally, the G-factor did not seem to be adequately defined as a domain-unspecific ASC construct: The loadings of GASC items on the G-factor ranged from $\lambda = .48$ to $.70$ in the model without school grades, but some domain-specific ASC items showed even higher loadings on the G-factor (e.g., items for physics ASC and chemistry ASC had loadings of $\lambda = .73$ and $\lambda = .83$, respectively).

When adding the school grades as outcome variables to the bifactor-ESEM representation, the factor loadings of the ASC items on the target S-factors and on the G-factor did not change in sign and were of similar size; however, the standard errors of the factor loadings (indicating precision of estimation) became considerably smaller. Therefore, all target loadings of the ASC items on their matching S-factors and on the G-factor became statistically significant, along with many

cross-loadings (Table 6). The problems regarding the definition of the G-factor also vanished, because none of the loadings of domain-specific ASC items on the G-factor exceeded the highest loading of a GASC item ($\lambda = .74$).

The S-factors for domain-specific ASCs demonstrated positive and statistically significant relations to the school grades of the matching domains (e.g., math ASC and math grade: $r = .38$; Table 4; see also Supplemental Table S5 in the online version of the journal). Across domains, most of the correlations between ASCs and school grades were not statistically significant; six out of 42 correlations were statistically significant but small in size ($r_s = -.10$ to $.16$). The S-factor for GASC was positively and substantially correlated to school grades in all subjects ($r_s = .32$ to $.43$). Similarly, the G-factor was found to share positive and substantial correlations with all school grades ($r_s = .14$ to $.41$).

First-Order Factor Model

The first-order factor model provided a good fit to the data (Table 2). All ASC factors were well-defined as indicated by substantially positive factor loadings ranging from $\lambda = .71$ for GASC to $\lambda = .92$ for chemistry ASC (Table 3). GASC showed statistically significant positive correlations with all domain-specific ASCs; these correlations ranged from $r = .38$ (for history ASC) to $r = .52$ (for math ASC; Table 5). Among different domain-specific ASCs, the correlations ranged from a small, yet statistically significant negative correlation between math and German ASCs ($r = -.12$; i.e., between a math-like and a verbal-like ASC) to a moderate and statistically significant positive correlation between physics and chemistry ASCs ($r = .60$; i.e., between two math-like ASCs).

When considering the relations to achievement, the highest correlations resulted between ASCs and school grades related to the same domain (e.g., math ASC and math grade: $r = .51$; Table 4; see also Supplemental Table S6 in the online version of the journal). GASC displayed positive and statistically significant correlations with all school grades ($r_s = .39$ to $.47$).

Discussion

Part 1 of this article provided an in-depth review of vital theoretical and methodological characteristics of five central structural models of ASC: the higher-order factor model, the Marsh/Shavelson model, the nested Marsh/Shavelson model, a bifactor representation based on ESEM, and the first-order factor model. In Part 2, we illustrated the application of these models using data from a large field trial with secondary school students from Germany. In the following sections, we discuss our empirical findings from a theoretical and methodological perspective. We then conclude with recommendations suggesting which research questions can be addressed by the different ASC models and by outlining advantages and shortcomings of each model.

Review of Empirical Findings

Multidimensionality

The assumption of multidimensionality was empirically supported in all models because all factors representing domain-specific ASCs were well-defined. In

particular, the domain-specific ASC items had substantial positive loadings on their respective ASC factors in all models. The domain specificity of ASC was further corroborated by the correlations between ASCs and school grades used as achievement indicators. The nested Marsh/Shavelson model, the bifactor-ESEM representation, and the first-order factor model allow for the examination of domain-specific outcome relations. In all these models, the strongest relations between ASC and school grades were observed in the same domain (i.e., school subject), whereas the correlations between ASCs and school grades of nonmatching domains were relatively lower. This indication of domain specificity was found irrespective of whether these relations involved first-order ASCs (in the first-order factor model) or residualized ASCs controlled for the G-factor (in the nested Marsh/Shavelson model and bifactor-ESEM representation).

Hierarchy

All models assuming a hierarchical ASC structure (i.e., the higher-order factor model, the Marsh/Shavelson model, the nested Marsh/Shavelson model, and the bifactor-ESEM representation) showed a satisfactory fit to the data. In addition, the constructs representing hierarchically superordinate ASC constructs (i.e., the HGASC in the higher-order factor model, the higher-order math and verbal ASCs in the Marsh/Shavelson model, and the G-factors in the nested Marsh/Shavelson model and the bifactor-ESEM representation) were well-defined by the respective subordinate domain-specific ASC factors.

In the higher-order factor model and in the bifactor-ESEM representation, the meaning of the HGASC and the G-factor, respectively, varies with the ASC domains included. This issue was clearly illustrated in our study that was based on a data set containing two verbal-like ASC measures (German, English), three math-like ASC measures (math, physics, chemistry), and two presumably mixed ASC measures (history, biology). This overrepresentation of math-like ASCs may explain the relatively higher factor loadings of the three math-like ASCs on the HGASC and relatively lower factor loadings of the two verbal-like ASCs on the HGASC in the higher-order factor model. A similar pattern of findings emerged in the bifactor-ESEM representation, in that the items for German and English ASCs showed low loadings on the G-factor despite defining strong S-factors. The items for math, physics, and chemistry ASCs, on the other hand, had substantial loadings both on their matching S-factors and on the G-factor. In contrast, in the nested Marsh/Shavelson model, the GASC items displayed the highest factor loadings on G-factor relative to the loadings of all domain-specific ASC items. This is a vital characteristic of the nested Marsh/Shavelson model: The G-factor retains its meaning irrespective of which and how many ASC domains are integrated (Eid et al., 2017).

The Meaning of GASC

Measures of GASC have been increasingly disregarded by ASC research given the strong domain specificity of ASC. However, the consideration and inclusion of GASC measures can be advantageous. First, GASC items are a direct and economical measure of students' ASC across a variety of domains or school subjects. The findings from this study corroborate this conclusion: In the higher-order

factor model, the GASC factor displayed the highest loading on the HGASC of all the first-order ASC factors. In the Marsh/Shavelson model, the GASC factor had similarly sized loadings on both the higher-order math and higher-order verbal ASCs. In the nested Marsh/Shavelson model, the GASC items showed higher loadings on the G-factor compared to the loadings of the domain-specific ASC items. In the bifactor-ESEM representation, the GASC items had substantial positive loadings on the G-factor. The first-order factor model treats GASC as another first-order factor that is positively correlated with the other domain-specific ASCs, and these correlations were of similar size for the various domain-specific ASCs.

The conclusion that GASC displays a cross-sectional representation of domain-specific ASCs is further supported by the inspection of the correlations to achievement. In the first-order factor model, the relations between the first-order GASC factor and the different school grades were all statistically significant, positive, and similar in size. The same pattern of relations could be observed for the G-factor that is defined by the GASC items in the nested Marsh/Shavelson model. Finally, in the bifactor-ESEM representation, the S-factor for GASC showed similarly sized relations to the different school grades. In sum, our findings consistently supported the capacity of GASC measures to capture how students perceive their school-related abilities in general.

We also observed a substantial proportion of residual variance of GASC in the higher-order factor model, the Marsh/Shavelson model, and the bifactor-ESEM representation. This residual variance suggests that GASC items capture some specific variance that is not shared with the domain-specific ASC items. This variance might refer to ASCs in noncore school subjects (e.g., physical education or arts), or might be related to secondary or side competences such as self-regulated learning or social/cooperative skills. Hence, a second advantage of GASC measures is that they seem to contain more information than what is reflected by domain-specific ASCs alone. Still, the residual variance of GASC might not only represent these secondary academic competences but also reflect students' differential weighting of school subjects or domains when responding to GASC items. In other words, students might have different school subjects or domains in mind when asked for their self-perceptions related to "all school subjects." The potential subjectivity of GASC items is a common characteristic of measures assessing general constructs not tied to specific domains. For example, Pavot and Diener (1993, p. 164) concluded that items assessing general life satisfaction (e.g., "In most ways my life is close to my ideal") have the advantage that individuals can "weight domains of their lives in terms of their own values, in arriving at a global judgment of life satisfaction." In a similar vein, GASC items have the advantage of allowing students to weight different school subjects according to what is important to them and to apply their own standards for success when being asked how they perceive their abilities related to "all school subjects." Studies explicitly asking students about the domain(s) they think of when responding to GASC items might help to further illuminate the meaning and subjectivity of GASC measures.

A third advantage of GASC measures is related to the finding that ASC measures and outcome variables (e.g., achievement, learning behavior) are the most

strongly related when both are located on the same level of hierarchy and when they address the same content (Swann et al., 2007). When the task is to explain and predict domain-unspecific outcomes (e.g., general academic achievement, persistence at school), GASC measures might be more useful than domain-specific ASC measures.

Conceptual Closeness of Domain-Specific ASCs

Apart from the higher-order factor model, all models presented here offer the possibility to inspect the conceptual closeness of domain-specific ASCs. The relevant models consistently demonstrated the separation between math and verbal ASCs as well as closer associations between different math-like ASCs and between different verbal-like ASCs, providing empirical support for a math-verbal continuum of ASCs. In the Marsh/Shavelson model, the higher-order math and verbal ASCs were nearly uncorrelated. In the nested Marsh/Shavelson model and the first-order factor model, math and German ASCs even showed a negative correlation. In addition, the correlations between math-like ASCs and verbal-like ASCs were considerably lower than the correlations within math-like ASCs or within verbal-like ASCs. In the nested Marsh/Shavelson model and in the first-order factor model, we found positive correlations between math, physics, and chemistry ASCs as math-like ASCs. Similarly, we found positive correlations between the verbal-like German and English ASCs. In the first-order factor model, the correlations between and among math-like and verbal-like ASCs partially reflect the common variance among domain-specific ASCs that can be attributed to a hierarchically superordinate construct of general ASC. In the nested Marsh/Shavelson model, however, these correlations are controlled for the G-factor. This might also explain the stronger (negative) correlation between math and German ASCs in the nested Marsh/Shavelson model as compared to the first-order factor model. In the bifactor-ESEM representation, the items measuring math ASC revealed negative cross-loadings on the German ASC factor, and vice versa. Furthermore, we found nonignorable negative cross-loadings between physics and German ASCs, and between chemistry and German ASCs, further illustrating the separation between math-like and verbal-like ASCs.⁴ Still, we found positive cross-loadings between math and physics ASCs and between German and English ASCs, illustrating the conceptual closeness among math-like and verbal-like ASCs, respectively.

History and biology ASCs were allocated in the center of the math-verbal continuum of ASC and thus were assumed to have both math-like and verbal-like characteristics. The findings of the present study lent support to the idea of considering biology ASC to be both math-like and verbal-like, as reflected by similar substantial loadings on the higher-order math and the higher-order verbal ASCs in the Marsh/Shavelson model. History ASC, however, seems to be more verbal-like than math-like, given its substantial loading on the higher-order verbal ASC but nonsubstantial loading on the higher-order math ASC. This conclusion was also supported by the findings from the bifactor-ESEM representation, which revealed positive cross-loadings between history and German ASCs. Moreover, in the nested Marsh/Shavelson model, we found a positive correlation between the S-factors for German ASC and history ASC but a negative correlation between

the S-factors for math ASC and history ASC. Finally, in the first-order factor model, the correlations between German or English ASCs and history ASC were higher than the correlation between math ASC and history ASC. Hence, findings from different structural models of ASC support the conclusion that history ASC is more verbal-like than math-like. This should be considered in future models on the structure of domain-specific ASCs; previous conceptualizations of history ASC as being both math-like and verbal-like should be revised.

Recommendations

“All models are approximations. Essentially, all models are wrong, but some are useful” (Box & Draper, 1987, p. 424). This common aphorism in statistics applies well to models of the ASC structure. All ASC models are associated with specific limitations that should be balanced against their advantages and usefulness when selecting the most appropriate model for a specific research question. We therefore conclude by offering researchers guidance in how to select an appropriate model contingent upon the specific research questions and study aims and by pointing out the main limitations associated with each ASC model.

Higher-Order Factor Model

Research questions. From a theoretical standpoint, the higher-order factor model is well suited to empirically testing the assumptions of Shavelson et al. (1976). From a methodological standpoint, the higher-order factor model is adequate to assess the variance shared across all ASC items included in a study as represented in the HGASC. In terms of application, the higher-order factor model might be useful in studies that aim to aggregate ASC scores across several domains in order to test the relation of this aggregated ASC score to outcome variables that are also measured on a general level, such as grade point average or general school satisfaction.

Limitations. First, in data sets that include a broad variety of domain-specific ASC measures like the one used in the present study, the higher-order factor model fits the data worse than alternative structural models of ASCs. That is, the higher-order factor model does not reflect the empirical relations among ASC measures as well as other models do. For instance, it does not incorporate the consistently observed differentiation between math and verbal ASCs. Second, the higher-order factor model suffers from the proportionality constraint. Hence, it is not well suited to study the relations between first-order ASCs (i.e., domain-specific ASCs or GASC) and outcome variables, because the ratio of variance attributable to a first-order ASC and to the higher-order factor is constrained to be the same across the ASC items. Third, the meaning of HGASC depends on the ASC domains under investigation, limiting the comparability of results across studies using data sets with different domain-specific ASC measures. If the goal for HGASC is to represent general ASC as accurately as possible, a high number of domains as well as a balance of math-like and verbal-like domains would be desirable. Finally, given its foundation on the ICM/CFA approach, the higher-order factor model does not take item cross-loadings

into account, although they are plausible in multidimensional constructs such as ASC.

Marsh/Shavelson Model

Research questions. From a theoretical standpoint, the Marsh/Shavelson model is useful for examining the location of domain-specific ASCs along a math-verbal continuum. From a methodological standpoint, the Marsh/Shavelson model is well suited to examining the common variance across math-like ASCs and across verbal-like ASCs, as reflected in the higher-order math and verbal ASCs. For instance, based on the Marsh/Shavelson model, researchers have examined which science domains are most math-like (Jansen et al., 2015). The model may also be useful for classifying students based on their preference for either the math or verbal domain, using their scores on the higher-order math versus higher-order verbal ASCs. When the need for parsimony only allows the inclusion of higher-order math and higher-order verbal ASCs instead of a variety of domain-specific ASCs, the higher-order ASCs can be useful for predicting students' educational choices. For example, one might predict that students' major choice at college or university in the science, technology, engineering, and mathematics domain is associated with higher scores on the higher-order math ASC than on the higher-order verbal ASC (Guo et al., 2015).

Limitations. First, as the Marsh/Shavelson model assumes two higher-order factors (i.e., a higher-order math ASC and a higher-order verbal ASC), it suffers from the proportionality constraint. This means that correlations among all first-order and higher-order ASCs and outcome variables cannot be examined with this model. Second, the meanings of the (math and verbal) higher-order constructs depend on the ASC domains under investigation, which limits the comparability of results across studies that include varying domain-specific ASC measures. A high number and balanced selection of math-like and verbal-like domains will result in a higher substantive validity of higher-order math and verbal ASCs. Third, the Marsh/Shavelson model cannot account for differences between ASCs within the math or verbal domains, as all math-like or verbal-like ASCs are assumed to commonly load on the higher-order math ASC or higher-order verbal ASC, respectively. Finally, the Marsh/Shavelson model relies on the ICM/CFA approach and therefore does not consider possible item cross-loadings across domain-specific ASCs.

Nested Marsh/Shavelson Model

Research questions. From a theoretical standpoint, the nested Marsh/Shavelson model is well suited to simultaneously testing the multidimensionality and hierarchy of ASC following the theoretical assumptions by Shavelson et al. (1976). From a methodological standpoint, as an incomplete bifactor model, the nested Marsh/Shavelson model is an interesting refinement and advancement of previous CFA models. The meaning of the G-factor is invariant even in cases of varying numbers and scope of the domain-specific ASC measures. Hence, this model may be particularly useful in integrative data analyses (e.g., across several

data sets based on independent samples; Curran, 2009), in longitudinal research (e.g., when different domain-specific ASC measures are considered over time to reflect the changing school curricula in different grade levels), or in studies that include an unbalanced number of domain-specific math-like and verbal-like ASCs. The nested Marsh/Shavelson model can further be used to examine the relations between all domain-specific ASCs and the G-factor and outcome variables (e.g., achievement). Finally, the nested Marsh/Shavelson model is ideally suited to depicting students' profiles of self-perceived strengths and weaknesses. In other words, students' scores on the S-factors for domain-specific ASCs represent their individual profiles of self-perceived strengths and weaknesses related to different domains, which are controlled for the G-factor (Schmidt et al., 2017). Hence, the nested Marsh/Shavelson model can be used to better understand the importance of individual profiles for students' developmental trajectories (e.g., course selection).

Limitations. First, the nested Marsh/Shavelson model does not include an S-factor for GASC, and thus cannot be used to study research questions that focus on GASC in terms of specific variance that is not shared across the domain-specific ASC measures included. Second, given its foundation on the ICM/CFA approach, the nested Marsh/Shavelson model does not take item cross-loadings into account, but these cross-loadings may explain at least some item variance over and above the target loadings on domain-specific ASCs and the G-factor.

Bifactor-ESEM Representation

Research questions. From a theoretical standpoint, the bifactor-ESEM representation simultaneously reflects the multidimensionality and hierarchy of ASC following the assumptions by Shavelson et al. (1976). With this model, researchers can test the relations between domain-specific ASCs, GASC, and the G-factor and outcome variables (e.g., achievement). From a methodological standpoint, it is a full bifactor model including an S-factor for GASC, which makes it possible to test whether GASC retains some specific variance over and above the G-factor and how this specific variance is related to outcome variables. Moreover, based on the ESEM approach, the bifactor-ESEM representation considers the inclusion of item cross-loadings, which are theoretically plausible in multidimensional constructs such as ASC. This model characteristic might be particularly useful for the purposes of scale development and evaluation. In fact, item cross-loadings reflect the extent to which ASC items measure conceptually close but distinct ASCs over and above the target ASCs. This empirical knowledge may help construct ASC scales that better measure a certain target construct.

Limitations. First, the bifactor-ESEM representation cannot probe for correlations among domain-specific ASCs (including GASC), as the corresponding factors are specified to be orthogonal. Second, the meaning of the G-factor reflecting general ASC may change contingent upon the number and selection of domain-specific ASCs considered. Hence, a high and balanced number of domain-specific math-like and verbal-like ASCs should be included in respective studies. Third,

in our study, some estimation problems concerning the size and statistical significance of factor loadings occurred in models without school grades as outcome variables. Hence, considerably more knowledge (e.g., on sample size, number of factor loadings, ratio of number of indicators per factor) is needed to identify the conditions for obtaining reliable parameter estimates in the bifactor-ESEM representation.

First-Order Factor Model

Research questions. From a theoretical standpoint, the first-order factor model depicts the multidimensionality of ASC, and therefore reflects one core assumption of the ASC structure as proposed by Shavelson et al. (1976). Moreover, this model enables researchers to examine the relations between domain-specific ASCs and GASC and outcome variables (e.g., achievement). Finally, the first-order factor model can probe for the conceptual closeness of ASCs within and across domains. From a methodological standpoint, it is surely the simplest model of the ASC structure and can thus be used as a starting point for data evaluation. Practically, many studies are based on the first-order model, particularly when they include only domain-specific ASC measures (i.e., no measure of GASC) and when they are not interested in possible higher-order constructs of ASC or in the structure of ASC. The first-order factor model is frequently used to examine relations between ASC and outcome variables, such as studies on the cross-sectional and longitudinal relations between ASC and achievement within one domain (e.g., Arens et al., 2017; Marsh et al., 2015) or across several domains (e.g., Möller et al., 2009, 2020; Weidinger et al., 2019).

Limitations. First, the first-order factor model does not consider the hierarchy of ASC; thus, it does not fully represent the theoretical assumptions of the nature of ASC (Shavelson et al., 1976). Consequently, the resulting domain-specific ASCs represent a blend of variance attributable to a hierarchically superordinate ASC construct and to domain-specific ASC facets. Second, given its foundation on the ICM/CFA approach, the first-order factor model does not take into account the theoretically plausible and empirically validated item cross-loadings of the multidimensional ASC construct.

General Remarks

An ideal approach to each ASC study would surely be to comparatively test and evaluate the different ASC models using multiple criteria including those applied in the present study. Researchers could thus learn about the consistency of empirical findings across different ASC models and get insight into how the results vary contingent upon which ASC model is applied. Still, this might not be always possible because the different ASC models rely on different conditions of use. The higher-order factor model, the Marsh/Shavelson model, and the bifactor-ESEM model need a broad array of domain-specific ASC measures as well as a GASC measure. Here, a balanced number of domain-specific math-like and verbal-like ASCs is recommended to prevent verbally or mathematically biased hierarchically superordinate constructs. Given the invariance of the G-factor in the nested

Marsh/Shavelson model, a minimal specification of the nested Marsh/Shavelson model is the assessment of one domain-specific ASC in addition to GASC items. On the other hand, the first-order factor model can be used even when only one domain-specific ASC is assessed. This might be the case in cross-sectional and longitudinal large-scale studies such as the Programme for International Student Assessment (PISA) or the Trends in International Mathematics and Science Study (TIMSS), which cover many student variables beyond ASC.

Concerning sample size requirements, the first-order factor model seems to have the least restrictive requirements. The simulation study by Wolf et al. (2013) implied that a sample of $N = 150$ students is sufficient to obtain reliable model parameters in a CFA model with three factors that are defined by substantial item indicators greater than $\lambda = .80$. Furthermore, the required sample size does not substantially increase when more than three well-defined factors are included. Hence, the required sample size depends on the measurement quality of the factor indicators (i.e., factor loadings of the item indicators), suggesting that the use of well-validated ASC instruments is important. Also, the required sample size varies contingent upon the number of indicators (e.g., items) per factor (Bollen, 1989), making it difficult to establish general rules on sample size requirements. With respect to more complex models involving at least one hierarchical superordinate ASC construct, it is even more challenging to postulate sample size requirements. Simulation studies would thus be helpful for the purpose of study planning (Muthén & Muthén, 2002) and can build on the empirical findings of the present study. For example, the simulation study by Morgan et al. (2015) suggested that a sample of $N = 200$ is required for analyzing higher-order factor models (with one higher-order and four first-order factors with two to three indicators per first-order factor) and bifactor-CFA models (with one G-factor and four S-factors with two to three indicators per S-factor).

Conclusion

Starting with the seminal theoretical model by Shavelson et al. (1976), researchers have been interested in investigating the relations among domain-specific ASCs, between domain-specific ASCs and GASC, and between ASC and outcome variables. In this context, researchers often face the question of which structural model of ASC should be used. We systematically compared and empirically illustrated those structural models that have been most often applied in past research or that have been recently established through methodological developments. To this end, we discussed in detail the inherent properties of the higher-order factor model, the Marsh/Shavelson model, the nested Marsh/Shavelson model, the bifactor-ESEM representation, and the first-order factor model. Moreover, we outlined how these models represent key theoretical assumptions and empirical findings concerning the structure of ASC, how the included factors can be interpreted, and how they can address substantial questions of ASC research and theory. Each ASC model has its advantages and limitations when it comes to answering different research questions. Hence, careful consideration is needed when selecting a specific ASC model. Other core constructs in education research (e.g., academic anxiety, academic interest) have theoretical underpinnings that are similar to ASC as they can also be conceptualized as

multidimensional and hierarchical in nature (Gogol et al., 2017). We therefore hope that our review provides helpful guidance not only to researchers choosing between structural models representing ASC but also to researchers focusing on other core constructs of education research and psychology.

ORCID iD

A. Katrin Arens  <https://orcid.org/0000-0003-2494-4199>

Notes

Preparation of this article was supported by a Heisenberg fellowship grant from the German Research Foundation to A. Katrin Arens (AR 877/3-1).

¹The hierarchy and multidimensionality assumption of ASC goes even further, as domain-specific ASCs can be separated into more specific facets. For instance, verbal ASC can be differentiated into reading, listening, and writing self-concepts (Arens & Jansen, 2016; Schmidt et al., 2017; Yeung et al., 2000). However, this review is limited to domain-specific ASCs.

²In this article, we consider only a higher-order factor model that includes general ASC as a first-order factor. This provides a clearer comparison to the other ASC models considered here, which all include a first-order factor for general ASC.

³Until 2012, these studies were referred to as “IQB National Assessment Studies.” They are now referred to as “IQB Trends in Student Achievement.”

⁴The differentiation between math-like and verbal-like ASCs has been theoretically explained by the internal/external frame of reference (I/E) model (Marsh, 1990a; Möller et al., 2009, 2020). In the supplements in the online version of the journal, we describe the assumptions of the I/E model in detail. In addition, we present the empirical findings obtained when estimating the I/E model with our data using the five different ASC models.

References

- 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., Marsh, H. W., Pekrun, R., Lichtenfeld, S., Murayama, K., & vom Hofe, R. (2017). Math self-concept, grades, and achievement test scores: Long-term reciprocal effects across five waves and three achievement tracks. *Journal of Educational Psychology, 109*(5), 621–634. <https://doi.org/10.1037/edu0000163>
- Arens, A. K., & Morin, A. J. S. (2017). 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>
- Asparouhov, T., & Muthén, B. O. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16*(3), 397–438. <https://doi.org/10.1080/10705510903008204>
- Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley.
- Box, G. E. P., & Draper, N. R. (1987). *Wiley series in probability and mathematical statistics. Empirical model-building and response surfaces*. John Wiley.

- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J.S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.
- 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>
- 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. (1996). *Measurement and instrumentation in psychology. Measuring self-concept across the life span: Issues and instrumentation*. American Psychological Association. <https://doi.org/10.1037/10197-000>
- 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
- Christensen, H., Mackinnon, A., Korten, A., & Jorm, A. F. (2001). The “common cause hypothesis” of cognitive aging: Evidence for not only a common factor but also specific associations of age with vision and grip strength in a cross-sectional analysis. *Psychology and Aging*, 16(4), 588–599. <https://doi.org/10.1037/0882-7974.16.4.588>
- Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6(4), 330–351.
- Curran, P. J. (2009). The seemingly quixotic pursuit of a cumulative psychological science: Introduction to the special issue. *Psychological Methods*, 14(2), 77–80. <https://doi.org/10.1037/a0015972>
- Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods*, 5(2), 155–174. <https://doi.org/10.1037/1082-989X.5.2.155>
- 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.
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford.

- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7(3), 286–299. <https://doi.org/10.1037/1040-3590.7.3.286>
- Diamantopoulos, A., & Siguaw, J. A. (2000). *Introducing LISREL*. Sage.
- Gignac, G. E. (2016). The higher-order model imposes a proportionality constraint: That is why the bifactor model tends to fit better. *Intelligence*, 55, 57–68. <https://doi.org/10.1016/j.intell.2016.01.006>
- 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>
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549–576. <https://doi.org/10.1146/annurev.psych.58.110405.085530>
- Graham, J. W., Hofer, S. M., Donaldson, S. I., MacKinnon, D. P., & Schafer, J. L. (1997). Analysis with missing data in prevention research. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 325–366). American Psychological Association.
- Guo, J., Parker, P. D., Marsh, H. W., & Morin, A. J. S. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology*, 51(8), 1163–1176. <https://doi.org/10.1037/a0039440>
- 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>
- Harter, S. (1982). The Perceived Competence Scale for Children. *Child Development*, 53(1), 87–97. <https://www.jstor.org/stable/1129640>
- Helmke, A., & van Aken, M. A. G. (1995). The causal ordering of academic achievement and self-concept of ability during elementary school: A longitudinal study. *Journal of Educational Psychology*, 87(4), 624–637. <https://doi.org/10.1037/0022-0663.87.4.624>
- Howard, J. L., Gagné, M., Morin, A. J. S., & Forest, J. (2018). Using bifactor exploratory structural equation modeling to test for a continuum structure of motivation. *Journal of Management*, 44(7), 2638–2664. <https://doi.org/10.1177/0149206316645653>
- Hox, J. J., Maas, C. J. M., & Brinkhuis, M. J. S. (2010). The effect of estimation method and sample size in multilevel structural equation modeling. *Statistica Neerlandica*, 64(2), 157–170. <https://doi.org/10.1111/j.1467-9574.2009.00445.x>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- 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>
- Kline, R. B. (2005). *Principles and practice of structural equation modeling*. Guilford.

- Litalien, D., Morin, A. J. S., Gagné, M., Vallerand, R. J., Losier, G. F., & Ryan, R. M. (2017). Evidence of a continuum structure of academic self-determination: A two-study test using a bifactor-ESEM representation of academic motivation. *Contemporary Educational Psychology, 51*, 67–82. <https://doi.org/10.1016/j.cedpsych.2017.06.010>
- Mai, Y., Zhang, Z., & Wen, Z. (2018). Comparing exploratory structural equation modeling and existing approaches for multiple regression with latent variables. *Structural Equation Modeling, 25*(5), 737–749. <https://doi.org/10.1080/10705511.2018.1444993>
- Marsh, H. W. (1987). The hierarchical structure of self-concept and the application of hierarchical confirmatory factor analysis. *Journal of Educational Measurement, 24*(1), 17–39. <https://doi.org/10.1111/j.1745-3984.1987.tb00259.x>
- Marsh, H. W. (1990a). 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. (1990b). 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. (2007). *Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. British Psychological Society.
- Marsh, H. W., Abduljabbar, A. S., Abu-Hilal, M. M., Morin, A. J. S., Abdelfattah, F. A., Leung, K. C., Xu, M., Nagengast, B., & Parker, P.D. (2013). Factorial, convergent, and discriminant validity of TIMSS math and science motivation measures: A comparison of Arab and Anglo-Saxon countries. *Journal of Educational Psychology, 105*(1), 108–128. <https://doi.org/10.1037/a0029907>
- 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. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective. Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*(2), 133–163. <https://doi.org/10.1111/j.1745-6916.2006.00010.x>
- Marsh, H. W., & Hattie, J. (1996). Theoretical perspectives on the structure of self-concept. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 38–90). John Wiley.
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to cutoff values for fit indexes and dangers in over-generalizing Hu & Bentler's (1999). *Structural Equation Modeling, 11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2
- Marsh, H. W., Liem, G. A. D., Martin, A. J., Morin, A. J. S., & Nagengast, B. (2011). Methodological measurement fruitfulness of exploratory structural equation model (ESEM): New approaches to key substantive issues in motivation and engagement. *Journal of Psychoeducational Assessment, 29*(4), 322–346. <https://doi.org/10.1177/0734282911423093>
- Marsh, H. W., Lüdtke, O., Nagengast, B., Trautwein, U., Abduljabbar, A. J. 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., 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). Guilford.
- Marsh, H. W., Muthén, B., Asparouhov, T., Lüdtke, O., Robitzsch, A., Morin, A. J. S., & Trautwein, U. (2009). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*, 16(3), 439–476. <https://doi.org/10.1080/10705510903008220>
- 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., 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>
- MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, 51, 201–226. <https://doi.org/10.1146/annurev.psych.51.1.201>
- McDonald, R. P. (1985). *Factor analysis and related methods*. Erlbaum.
- 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., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). A meta-analysis of relations between achievement and self-concept. *Review of Educational Research*, 90, 376–419. <https://doi.org/10.3102/0034654320919354>
- Morgan, G. B., Hodge, K. J., Wells, K. E., & Watkins, M. W. (2015). Are fit indices biased in favor of bi-factor models in cognitive ability research? A comparison of fit in correlated factors, higher-order, and bi-factor models via Monte Carlo simulations. *Journal of Intelligence*, 3(1), 2–20. <https://doi.org/10.3390/jintelligence3010002>
- Morin, A. J. S., Arens, A. K., & Marsh, H. W. (2016). A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling*, 23(1), 116–139. <https://doi.org/10.1080/10705511.2014.961800>
- Morin, A. J. S., Marsh, H. W., & Nagengast, B. (2013). Exploratory structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 395–436). Information Age.
- Muthén, L. K., & Muthén, B. O. (1998–2019). *Mplus user's guide* (8th ed.). Muthén & Muthén.
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 9(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8
- Pant, H. A., Stanat, P., Schroeders, U., Roppelt, A., Siegle, T., & Pöhlmann, C. (2013). *The IQB National Assessment Study 2012: Competencies in mathematics and the sciences at the end of secondary level I*. Waxmann.

- Pavot, W., & Diener, E. (1993). Review of the Satisfaction With Life Scale. *Psychological Assessment*, 5(2), 164–172. <https://doi.org/10.1037/1040-3590.5.2.164>
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667–696. <https://doi.org/10.1080/00273171.2012.715555>
- Reise, S. P., Moore, T. M., & Maydeu-Olivares, A. (2011). Targeted bifactor rotations and assessing the impact of model violations on the parameters of unidimensional and bifactor models. *Educational and Psychological Measurement*, 71(4), 684–711. <https://doi.org/10.1177/0013164410378690>
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177. <https://doi.org/10.1037/1082-989X.7.2.147>
- 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>
- Schmiedek, F., & Li, S.-C. (2004). Toward an alternative representation for disentangling age-associated differences in general and specific cognitive abilities. *Psychology and Aging*, 19(1), 40–56. <https://doi.org/10.1037/0882-7974.19.1.40>
- 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>
- Stapleton, L. M. (2006). An assessment of practical solutions for structural equation modeling with complex sample data. *Structural Equation Modeling*, 13(1), 28–58. https://doi.org/10.1207/s15328007sem1301_2
- Swann, W. B., Chang-Schneider, C., & Larsen McClarty, K. (2007). Do people's self-views matter? Self-concept and self-esteem in everyday life. *American Psychologist*, 62(2), 84–94. <https://doi.org/10.1037/0003-066X.62.2.84>
- 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
- Weidinger, A. F., Steinmayr, R., & Spinath, B. (2019). Ability self-concept formation in elementary school: No dimensional comparison effects across time. *Developmental Psychology*, 55(5), 1005–1018. <https://doi.org/10.1037/dev0000695>
- West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). Guilford.
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 73(6), 913–934. <https://doi.org/10.1177/0013164413495237>
- Wolff, F., Nagy, G., Retelsdorf, J., Helm, F., Köller, O., & Möller, J. (2019). The 2I/E model: Integrating temporal comparisons into the internal/external frame of reference model. *Journal of Educational Psychology*, 111(7), 1131–1161. <https://doi.org/10.1037/edu0000319>
- Yeung, A. S., 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>

Authors

A. KATRIN ARENS, PhD, holds a Heisenberg postdoctoral position funded by the German Research Foundation (DFG) at the DIPF/Leibniz Institute for Research and Information in Education, Department on Research on Education and Human Development and Centre for Research on Individual Development and Adaptive Education of Children at Risk (IDeA), Rostocker Str. 6, D-60323 Frankfurt am Main, Germany; email: *arens@dipf.de*. Her research interests are motivation in education with an emphasis on self-concept and expectancy-value theory.

MALTE JANSEN, PhD, is head of the research data center at the Institute for Educational Quality Development at the Humboldt University of Berlin, Unter den Linden 6, 10099 Berlin, Germany; email: *malte.jansen@iqb.hu-berlin.de*. The research data center is also part of the center for international student assessment (ZIB). Malte Jansen's main research interests are student motivation, social and dimensional comparisons and social networks.

FRANZIS PRECKEL, PhD, is a full professor and head of the Chair of Giftedness Research and Education at the Department of Psychology, University of Trier, Universitätsring 15, D-54296 Trier, Germany; email: *preckel@uni-trier.de*. Her main research interests are intelligence, giftedness, factors influencing academic development and achievement, and psychological assessment.

ISABELLE SCHMIDT, PhD, is a senior researcher at GESIS-Leibniz Institute for the Social Sciences, B2 1, 68159 Mannheim, Germany; email: *isabelle.schmidt@gesis.org*. She is responsible for the publication of the open access repository for measurement instruments ZIS (Collection of Items and Scales for the Social Sciences). Her current research topics are psychological diagnostics and test methods, statistical modeling and data analysis of longitudinal data, and self-concept research in educational contexts.

MARTIN BRUNNER, PhD, is a full professor and head of the Chair of Quantitative Methods in Educational Sciences at the Department of Educational Sciences, University of Potsdam, Karl-Liebknecht-Straße 24–25, Potsdam, Brandenburg 14476, Germany; email: *mabrunner@uni-potsdam.de*. His main research interests are quantitative methods, educational assessment, and cognitive and noncognitive educational outcomes.