



A Meta-analysis of the Longitudinal Relationship Between Academic Self-Concept and Academic Achievement

Huimin Wu¹ · Yiqun Guo² · Yingkai Yang¹ · Le Zhao³ · Cheng Guo¹ 

Accepted: 26 January 2021 / Published online: 23 March 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

The reciprocal relationship between academic self-concept (ASC) and academic achievement has been documented in multiple studies. However, this relationship has not been investigated fully from a developmental perspective. In the present meta-analysis, 240 effect sizes were aggregated from 68 longitudinal studies to examine the longitudinal relationship between ASC and achievement. The results found that achievement significantly predicted ASC ($\beta = 0.16$, $p < 0.01$) and vice-versa ($\beta = 0.08$, $p < 0.01$) after controlling for the initial level of outcome variables, which provided further evidence for the reciprocal effects model (REM). Moderator analyses found that the effect of achievement on ASC was significantly moderated by student age, whereas the effect of ASC on achievement was significantly moderated by student age, achievement level, and types of achievement measurement. Combining the significant moderating effect of age on the paths leading from ASC to achievement and from achievement to ASC, the relationship between ASC and achievement was found to demonstrate a trend from a strong skill-development effect to a pronounced reciprocal effect with age within the framework of the REM.

Keywords Academic self-concept · Academic achievement · Meta-analysis · Reciprocal effects model

The question of how to promote students' academic achievement and educational attainment is of utmost concern in educational settings (Denissen et al., 2007; Marsh & Martin, 2011).

✉ Cheng Guo
guochen@swu.edu.cn

¹ Faculty of Psychology, Southwest University, No 2, Tiansheng Road, Beibei District, Chongqing, China

² School of Innovation and Entrepreneurship Education, Chongqing University of Posts and Telecommunications, Chongqing, China

³ School of Education, Beijing Normal University, Zhuhai, China

Academic self-concept (ASC) is considered a pivotal factor that influences a number of educational variables such as academic achievement (Valentine et al., 2004), achievement goals (Niepel et al., 2014), academic interests (Marsh et al., 2005), educational attainment level (Marsh & O'Mara, 2008), and course selections (Marsh & Yeung, 1997).

Over the past few decades, extensive research has been carried out to investigate for the relationship between ASC and achievement. Since the 1970s, three models have been proposed to describe the relationship: the self-enhancement model, the skill-development model, and the reciprocal effects model (REM) (Calsyn & Kenny, 1977; Marsh & Craven, 2006). According to the self-enhancement model, ASC is the determinant of academic achievement, not vice-versa. The skill-development model posits that academic achievement predicts ASC but that ASC does not affect achievement (Calsyn & Kenny, 1977). REM, another model that reconciles the two aforementioned either-or views, postulates that early ASC influences later academic achievement and prior achievement leads to gains in subsequent ASC within domains over time (Marsh & Craven, 2006; Marsh & Martin, 2011). Although previous longitudinal studies supported the REM (e.g., Chen et al., 2013; Möller et al., 2014; Niepel et al., 2014; Sewasew & Schroeders, 2019), this relationship has not been investigated fully from a developmental perspective.

In the present study, we conducted a meta-analysis to systematically integrate previous longitudinal studies on the relationship between ASC and achievement, aiming to provide further evidence for REM and examine the development of the relationship between ASC and achievement from childhood to adolescence. In addition to student age, we investigated other theoretical and methodological moderators that might influence the relationship between ASC and achievement, including student achievement level, types of achievement measurement, domain of measurement, match between domain specificity of ASC and achievement, time lags between measurement waves, variable types, sample types, and publication year.

The Reciprocal Effects Model

Shavelson et al. (1976) broadly defined self-concept as “a person’s self-beliefs formed through experience with and interpretations of his or her environment (p. 411)”. In an academic context, ASC has received extensive attention because of its direct relevance to educational settings. It pertains to one’s beliefs about ability in academic domain (Susperreguy et al., 2017) and is typically measured by “being” and “feeling” questions such as “I get good marks in (a subject)” or “I learn things quickly in (a subject)” (see Marsh, 1990; Wigfield et al., 1997). This is distinguished from academic self-efficacy which refers to the learners’ convictions that they could perform well in academic tasks to achieve their goals (Bandura, 2010).

While ASC and academic self-efficacy share some commonalities (e.g., being self-related beliefs, predictions of educational outcomes and emotion), three key differences should be kept in mind: (a) the central element: ASC mainly taps the individuals’ perceived academic competence based on past accomplishments and circumstances, whereas academic self-efficacy refers to people’s perceived confidence in successfully coping with different tasks in academic situations (Bong & Skaalvik, 2003; Marsh et al., 2019; Pietsch et al., 2003); (b) composition: more recent studies have shown that ASC is separated into cognitive and affective components (e.g., Arens et al., 2014; Arens & Hasselhorn, 2015; Kadir et al., 2017; Pinxten et al., 2014). The cognitive component is concerned with students’ sense of competence or the judgement of their own academic ability in a specific subject, as measured

by items such as “Are you good at (a subject)” or “Do you get good marks in (a subject)” (Bong & Skaalvik, 2003; Kadir et al., 2017). The affective component taps individuals’ affective-motivational reactions toward the domain by asking the items such as “Are you interested in (a subject)” or “Do you enjoy in (a subject)” (Bong & Skaalvik, 2003; Kadir et al., 2017). However, academic self-efficacy “exclusively taps the cognitive aspect of student’ self-perceptions” (Bong & Skaalvik, 2003, p. 13); (c) the nature of evaluation: the formation of ASC is most heavily influenced by social and dimensional comparison, whereas the level of academic self-efficacy primarily depends on individuals’ previous mastery experience in similar tasks (Bong & Skaalvik, 2003; Pietsch et al., 2003). Given the distinctions between the two constructs, our study focuses on the relationship between ASC and achievement.

The nature of the relationship between ASC and achievement is one of the “thorniest issues” in the field of educational psychology (Pajares & Schunk, 2001) and remains to be resolved. As Marsh and Martin (2011) and Shavelson et al. (1976) mentioned, achievement is the main source of the formation of ASC, so there is no need to argue the effect of achievement on ASC. The reciprocal internal/external frame of reference model also proposes that individuals develop and shape their ASC by social, dimensional, and temporal comparisons (Sewasew & Schroeders, 2019). That is, students perceive their domain-specific competence by comparing their achievement in a particular domain (e.g., a school subject) with the achievement of their classmates in the same domain (social comparison; Festinger, 1954), with their achievement in other different subjects (dimensional comparison; Möller & Marsh, 2013), and with their own prior achievement in the same domain (temporal comparison; Albert, 1977).

However, does a link from prior ASC to subsequent achievement in fact exist? Self-belief is an important driver of people’s behaviors and thoughts (Graham & Weiner, 1996). It is partly on the basis of beliefs that people choose what challenges to take on, how much effort to expend in the endeavor, and how long to persist in the face of obstacles (Bandura, 2010; Graham & Weiner, 1996). In addition, it has been suggested that students who have a positive view of themselves might strive to align their behavior and performance with their self-image (Sherman, 2013; Sherman & Cohen, 2006). It appears that students with a positive ASC are more motivated to self-regulation (e.g., adopting effective learning strategies, putting in efforts and persisting) to protect the consistency between self-concept and behaviors, and thus achieve academic success in a domain (Marsh, 1990; Marsh & Martin, 2011; Wigfield & Eccles, 2000). Hence, the process of linking ASC and achievement is likely to be reciprocal in theory: success in a domain enhances students’ perception of competence and gives them an intrinsic pleasure in the process of mastering the subject. As students work and become more skillful in a domain, the sense of competence for performing well would be maintained. In turn, their perceived high competence motivates them to invest more effort to achieve a cascade of success in future (e.g., Harter, 1978; Marsh & Martin, 2011).

To date, the reciprocal effect of ASC and achievement has been documented in some studies. Based on data from a large project in Germany ($N = 1, 508$), Möller et al. (2011) found that ASC and achievement were reciprocally related in both mathematical and verbal domains. Using two independent samples from an ongoing longitudinal study focusing on students’ self-concept development, Niepel et al. (2014) replicated the REM within the same domain. Sewasew and Schroeders (2019) assessed 16,216 students comprising a nationally representative sample from diverse ethnic and socioeconomic backgrounds and revealed a positive reciprocal relationship between ASC and achievement in the verbal and mathematical domains. Moreover, Valentine et al. (2004) identified 56 longitudinal studies exploring the relationship between self-beliefs (i.e., self-concept, self-esteem, self-efficacy) and

achievement, supporting the REM. Based on path analyses, Huang's (2011) study, in which 32 longitudinal studies investigating the relationship between self-concept (i.e., self-concept, self-esteem) and achievement were identified, also supported the REM.

However, although it is generally accepted that ASC and achievement are reciprocally related, some results in the extant literature did not support the REM. Marsh (1990) evaluated the causal order of general ASC and achievement using a four-wave design (i.e., early 10th grade, late 11th grade, late 12th grade, and 1 year after normal high school graduation). Results revealed that prior ASC affected subsequent grades but that prior grades had no effect on subsequent ASC, thereby supporting the self-enhancement model. Viljaranta et al. (2014) assessed students' ASC and achievement (i.e., teacher-rating and school grade) when they were in Grades 1, 2, 4, and 7, and the results supported the skill-development model. These mixed results might be explained by the diversity of participants, measurement instruments, or methodological characteristics across studies.

Relationship Between ASC and Achievement from a Developmental Perspective

Previous studies have provided support for the generalizability of REM across age (e.g., Guay et al., 2003; Helmke & van Aken, 1995; Marsh et al., 1999). For instance, Guay et al. (2003) evaluated general ASC and achievement in elementary students ($Mean_{age} = 9$ years old) on three occasions with 1-year interval and found that the REM was the best fitting model for all waves and age cohorts. Skaalvik and Valås (1999) assessed mathematical and verbal achievement and the corresponding self-concept twice for three cohorts of children (i.e., Grades 3, 6, and 8 at the beginning of the study). Their results indicated that the relationship between ASC and achievement was consistent with the skill-development model in childhood and adolescence, and no developmental effects were found.

However, some researchers have suggested a developmental perspective in which the nature of the relationship between ASC and achievement might demonstrate a trend from the skill-development model to the reciprocal model with age (Skaalvik & Hagtvet, 1990; Weidinger et al., 2018; Wigfield & Karpathian, 1991). In preschool and early elementary school, children tend to overestimate their academic competence and have an extremely positive ASC that is typically less relevant to external indicators such as grades and standardized tests (Marsh, 1990; Spinath & Spinath, 2005). Helmke and van Aken (1995) concluded that during elementary school, ASC is dominated mainly by cumulative achievement-related success and failure and has a minimally significant impact on subsequent grade or test performance. As students grow older, their ASC becomes increasingly realistic and differentiated, and depends more strongly on external achievement criteria (Marsh & Craven, 2006; Wigfield et al., 1997).

Some studies have lent support for the developmental hypothesis. Skaalvik and Hagtvet (1990) conducted a study with two cohorts of Norwegian school children who were assessed on two occasions on self-concept of ability and teacher-rating. Results showed a skill-development model in Grades 3 and 4 but a reciprocal model in Grades 6 and 7.

Weidinger et al. (2018) examined this developmental perspective using 542 children (Time [T]1: $Mean_{age} = 7.95$ years old) whose mathematics competence beliefs and grades were evaluated on seven occasions, with an interval of four months between each wave. The results found that mathematics competence beliefs declined over time and the direction of the

relationship between mathematics competence beliefs and mathematics grades changed over time. That was, at the end of second and the beginning of third grade, the skill-development model predominated over the self-enhancement model; at the end of elementary school, the REM was predominant.

Chen et al. (2013) assessed students' self-concept in the mathematical and verbal domains using a multiple cohort (i.e., Grades 5 and 10 at the beginning of the study), multiple occasions (two occasions with a 1-year interval) design. Results showed that for elementary school students in childhood, the skill-development model was found in the verbal domain but the REM was found in the mathematical domain; for high school students in adolescence, ASC and achievement were reciprocally related in both domains. In addition, the effect of achievement decreased with age, whereas the effect of ASC increased with age, indicating a developmental trend. An overview of longitudinal studies that could suggest the developmental effect was presented in Table 1.

To the best of our knowledge, most of the current studies have examined the development of the relationship between ASC and achievement in childhood or adolescence (e.g., Helmke & van Aken, 1995; Sewasew & Schroeders, 2019; Viljaranta et al., 2014; Weidinger et al., 2018), and few studies directly explored this development from childhood to adolescence. Moreover, in their meta-analysis, Valentine et al. (2004) explored only the moderating role of students age upon the effect of self-beliefs on achievement and found that age was not a significant moderator. This focus may have been insufficient for an examination of the development of the relationship between self-beliefs and achievement. Therefore, this relationship has not been investigated fully from a developmental perspective. The present meta-analysis tests whether the nature of the relationship between ASC and achievement changes with age by simultaneously examining both the moderating effect of participants' age on the paths leading from ASC to achievement and on that from achievement to ASC.

Potential Moderators of Relationship Between ASC and Achievement

In addition to students' age, some studies have shown that student achievement level affects the relationship between ASC and achievement (Gottfried et al., 2007; Prast et al., 2018). Academic experience varies from student to student: some students often experience academic success while some students often experience failure (Prast et al., 2018). Gottfried et al. (2007) showed that initial academic experience in school had long-term implications for students' subsequent motivation and achievement. For instance, compared with typical-achieving students, low-achieving students experience less accomplishment, less positive reinforcement from others, less academic motivation, and more academic anxiety, which may weaken the association between ASC and achievement (Hampton & Mason, 2003; Prast et al., 2018). However, some longitudinal research demonstrated that the relationship between ASC and achievement was similar across students with different achievement levels (Gorges et al., 2018; Möller et al., 2014; Seaton et al., 2015). Hence, it is unclear whether student achievement levels affect the relationship between ASC and achievement.

The differences in the types of achievement measurement across studies might also account for the inconsistency among conclusions. First, Meyer et al. (2019) argued that the extent of the association between ASC and achievement differs according to types of achievement measurement. Achievement is typically measured by school grades, teacher-rating, or standardized tests, which separately capture different aspects of achievement (Pinxten et al., 2010;

Table 1 Overview of longitudinal studies investigating the development change of the relationship between ASC and achievement (ordered according to the publication year)

Study	<i>N</i>	Age	Domain	Model
Skaalvik & Hagtvet, 1990	635	Cohort 1: 3rd grade → 4th grade (<i>Mean</i> _{age} = 9.25); Cohort 2: 7th grade → 8th grade (<i>Mean</i> _{age} = 12.25)	General	Cohort 1: SDT Cohort 2: REM
Helmke & van Aken, 1995	697	T1 → T2: 2nd grade → 3rd grade; T2 → T3: 3rd grade → 4th grade	Math	T1 → T2: SDT T2 → T3: SDT
Chapman & Tunmer, 1997	112	T1 → T2: 1st grade → 2nd grade (<i>Mean</i> _{age} = 5.61); T2 → T3: 2nd grade → 3rd grade (<i>Mean</i> _{age} = 6.61)	Verbal	T1 → T2: NULL T2 → T3: SDT
Skaalvik & Valås, 1999	1005	Cohort 1: 3rd grade → 4th grade; Cohort 2: 6th grade → 7th grade; Cohort 3: 8th grade → 9th grade	Math; Verbal	Support SDT in both domains in all cohorts
Guay et al., 2003	385	T1 → T2: 2nd, 3rd, 4th grade → 3rd, 4th, 5th grade (<i>Mean</i> _{age} = 9.5); T2 → T3: 3rd, 4th, 5th grade → 4th, 5th, 6th grade (<i>Mean</i> _{age} = 10.5)	General	T1 → T2: REM T2 → T3: REM
Pinxten et al., 2010	1753	T1 → T2: 7th grade → 8th grade; T2 → T3: 8th grade → 10th grade; T3 → T4: 10th grade → 12th grade	General	T1 → T2: REM T2 → T3: REM T3 → T4: REM
Chen et al., 2013	782	Cohort 1: 5th grade → 6th grade; Cohort 2: 10th grade → 11th grade	Math; Verbal	Cohort 1: SDT in verbal but REM in math Cohort 2: REM in both domains
Viljaranta et al., 2014	178	T1 → T2: 1st grade → 2nd grade (<i>Mean</i> _{age} = 8); T2 → T3: 2nd grade → 4th grade (<i>Mean</i> _{age} = 9.5); T3 → T4: 4th grade → 7th grade (<i>Mean</i> _{age} = 12)	Math; Verbal	Math: support REM in all waves Verbal: support REM at T1 → T2 and T3 → T4
Weidinger et al., 2018	542	T1 → T5: at the end of 2nd and the beginning of 4th grade (<i>Mean</i> _{age} = 8.17); T5 → T7: at the beginning of 4th and the end of 4th grade (<i>Mean</i> _{age} = 9.61)	Math	T1 → T5: SDT T5 → T7: REM

Only studies with three (or more) waves and/or more than one cohort are listed here. *SDT* skill-development model; *REM* reciprocal effects model; NULL indicates that ASC and achievement were not significantly correlated; *T* measurement occasion

Valentine et al., 2004). Because ASC possesses motivational properties, when achievement is measured in a way that is more strongly related to motivation, it is also more strongly related to the ASC (Wylie, 1974). Standardized tests are typically conducted by research institutions or teams who commonly cannot announce test results or students' relative level; thus, they are low stake and not of personal significance for the students themselves (Pinxten et al., 2010). However, school grades and teacher-rating often take additional information about students (e.g., academic effort or diligence) into account when rating students' performance (Helmke & van Aken, 1995; Nagengast et al., 2013). Hence, it is hypothesized that achievement based on grades or teacher-rating is more closely related with ASC than studies using standard tests.

In addition, some studies have revealed that the relationship between ASC and achievement is remarkably domain specific (Chen et al., 2013; Liem et al., 2015). Chen et al. (2013) and Möller et al. (2014) found that REM was supported in the mathematical but not in the verbal domain. The domain specificity might be related to the characteristics of different subjects. The verbal domain that typically contains a variety of learning materials (e.g., poems, biographies) and classroom activities (e.g., grammar training, reading), is more broadly defined. In contrast, the topics and formats covered in the mathematical domain are more narrowly defined, primarily including problem solving exercises and nonverbal knowledge (Dweck, 1986; Götz et al., 2010; Stodolsky & Grossman, 1995), so students' perceived competence and achievement are based on a more limited range of information sources. Therefore, it is speculated that subject domain could moderate the relationship between ASC and achievement.

Furthermore, Marsh's early research (1992) found that the associations between achievement and ASC in corresponding subjects were more highly correlated ($r_{\text{mean}} = 0.57$) than those in noncorresponding subjects ($r_{\text{mean}} = 0.33$). According to the specificity-matching principle, a domain specific self-concept shows more predictive power for outcomes that belong to the same degree of specificity; therefore, ASC is expected to show the greatest predictive power for achievement that exhibits the same degree of specificity (Swann Jr et al., 2007). Thus, we aimed to examine whether studies with domain matching measures show a higher correlation than those studies without matching measures.

In addition to the moderators proposed above, we also attempted to examine some potential methodological moderators including time lags between measurement waves, variable types, and sample types. Collins and Graham (2002) stated that it is important to highlight the effect of time lags on effect sizes when including longitudinal studies in meta-analysis. Theoretically, a minimum amount of time is required to detect prospective effects between ASC and achievement. The effect sizes are zero when the time interval is zero; the effect sizes increase and reach a maximum when the time intervals lengthen to a specific time interval, and then the effect sizes should decrease to close to zero after long time interval (Cole & Maxwell, 2003; Taris & Kompier, 2014). A previous meta-analysis by Valentine et al. (2004) found that the relationship between self-belief and achievement varied with the time intervals, and that effect sizes based on longer time intervals (> 18 months) were larger than those based on shorter time intervals (< 6 months). We therefore anticipated that time lags between measurement waves would moderate the relationship between ASC and achievement.

Leonhart et al. (2008) showed that compared with variables calculated at the manifest level, latent constructs could yield a more valid and potentially higher estimates. Since latent variables cannot be measured directly, they are typically operationalized by two or more manifest variables through structural equation modeling (SEM). With reference to the assumptions of Classical Test Theory, a combination of answers to two or more questions might provide a more accurate estimate of the latent variable behind these manifest variables (Hair et al., 1998; Loehlin, 2004). In addition, the use of latent variables and SEM could correct for unreliability of scales and estimate the measurement error variance directly (Leonhart et al., 2008; Loehlin, 2004). Given these considerations, the latent variables analysis may yield a less biased, more valid estimation of true values and a better representation of theoretical constructs. Therefore, whether the use of latent variables affects the relationship between ASC and achievement is an issue of interest to us.

Moreover, most of the longitudinal studies included in our current meta-analysis used either representative sampling or convenience sampling. *Convenience sampling* is a non-

probabilistic sampling or non-random sampling technique in which members of the target population who usually meet a series of realistic criteria, such as easy recruitment for the study, the willingness to participate, or geographical proximity (Etikan et al., 2016; Sedgwick, 2013). However, there are clear disadvantages to convenience sampling: the foremost being the low external validity due to sample bias. Moreover, the variability and bias cannot be measured or controlled (Acharya et al., 2013). In contrast, *representative sampling* usually utilizes random selection procedures and weighting techniques to maximize the representativeness of a sample (Zelinski et al., 2001). Two different sampling techniques might result in differences in the quality of the data collected. Therefore, we attempted to examine that whether the effect sizes would vary with sample types.

Finally, the latest publication year of the studies included in the Valentine et al.'s meta-analysis (2004) is 2001. However, 49 studies included in the current meta-analysis were published after 2001. Hence, we used the publication year as a moderator to investigate whether studies after 2001 have supported the REM.

The Present Meta-analysis

The first aim of the present meta-analysis was to extend on previous work by systematically integrating longitudinal studies on the relationship between ASC and achievement, and to provide further evidence for the REM. The one-stage meta-analytic SEM (OSMASEM) analysis that combines the advantages of meta-analysis and SEM was performed to examine whether the REM was established.

In addition, given that the relationship between ASC and achievement has not been examined fully from a developmental perspective, a secondary aim of our study was to test the development of the relationship between ASC and achievement from childhood to adolescence by simultaneously examining both the moderating effect of participant age on the path leading from ASC to achievement and on that from achievement to ASC.

Finally, previous studies have not yielded consistent conclusions on the moderating effect of some variables. We systematically investigated some theoretical and methodological moderators that might influence the relationship between ASC and achievement, including student achievement level, types of achievement measurement, domain of measurement, match between domain specificity of ASC and achievement, time lags between measurement waves, variable types, sample types, and publication year.

Method

Study Selection and Eligibility Criteria

Literature Search Relevant studies were identified by an exhaustive search that proceeded in three ways. First, we searched reviews and research papers on electronic databases, including PsycINFO, PubMed, Web of Science, ERIC and Google Scholar using combinations of the keywords (a) (“self-concept” OR “academic perception” OR “academic construction” OR “academic perspective” OR “academic structure” OR “expectancy-value belief” OR “competence belief”), paired with (b) (achievement OR performance OR tests OR grade OR examination OR “grade point average”), and (c) (longitudinal OR causal OR prospective OR

retrospective OR “risk factor” OR vulnerability OR antecedent OR cross-lagged OR predict OR “follow-up”). A full-text search strategy was used to avoid omitting eligible studies to the greatest extent possible. For all articles considered, we reviewed the abstract and examined the full texts whenever a paper might include a relevant effect. To reduce the publication bias, we also searched for as many unpublished dissertations and other unpublished studies as possible on ProQuest Dissertations and Theses and Google Scholar. The final research was conducted in November 2020. Second, if a study was identified as an eligible article, its reference list was further reviewed to examine if there were additional eligible studies. Third, the reference sections of previous two meta-analyses (i.e., Valentine et al., 2004; Huang, 2011) on the relationship between self-concept and achievement were screened. A total of 1725 abstracts were scanned for potential inclusion in this review (see Fig. 1 for PRISMA flow diagram).

Eligibility Criteria Studies that fulfilled the following criteria were included in the current meta-analysis: (a) ASC and achievement were assessed with continuous and explicit measures. (b) To avoid confusion between ASC and other related self-beliefs, only studies that conformed with the definition of ASC were included, while those measuring academic self-efficacy and self-esteem were excluded. If a study assessed expectancy value belief but ASC was used to operationalize the expectancy component, this study was included. (c) The study used a longitudinal design. (d) For OSMASEM analysis, only studies in which ASC and achievement were measured simultaneously at two time points were included. (e) When separately analyzing two crossed-lag paths, studies in which at least one of the constructs (e.g., ASC or achievement) was measured at two time points were included. (f) Sufficient information was provided to calculate effect sizes. Only studies published in English were included owing to the language restriction of researchers. There were no restrictions on publication date of studies.

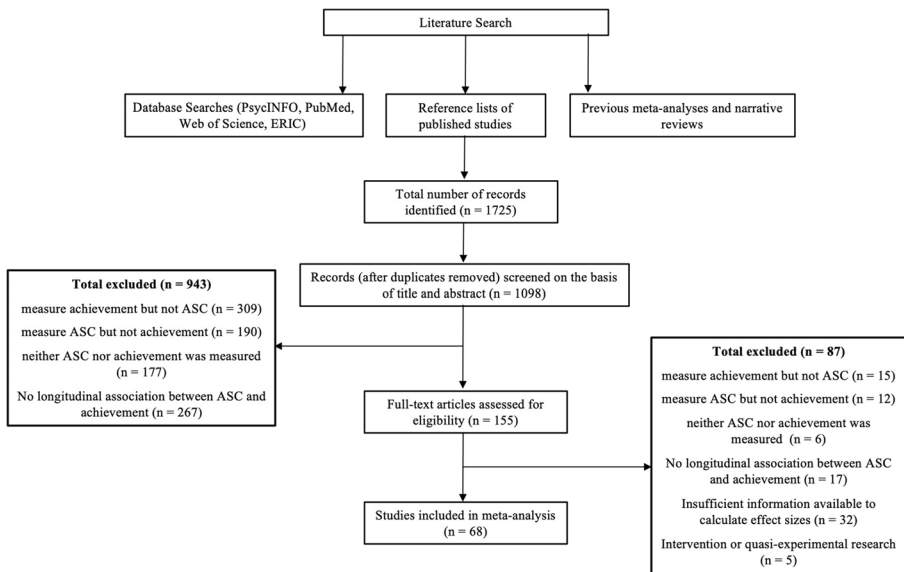


Fig. 1 PRISMA flow diagram illustrating the process of identifying eligible studies

Study Selection A total of 155 articles were assessed in full text, 68 of which met the eligibility criteria, including 61 peer-reviewed papers, 5 unpublished dissertations, and 2 conference paper. The authors of the remaining 87 studies that met the following criteria were contacted: (a) ASC and achievement were taken as the central constructs, and (b) most of the data for calculating the effect size were provided (e.g., only one correlation coefficient was missing). We eventually contacted the authors of 19 studies via e-mail but did not receive any replies. See Fig. 1 for specific reasons for study exclusion at each stage of the search process.

Independence of Effect Sizes Since large-scale longitudinal studies require enormous resources in terms of humans, energy, and finance, some studies that met the eligibility criteria may use overlapping samples. If several papers were based on the same sample, we only included one study that met the following criteria (in order of importance): (a) the study was most directly related to the current research. For example, if one study provided data on both ASC and achievement but another included data on only one of these concepts, the study with more information was selected; (b) the study contained the most comprehensive coding information and complete statistical data; (c) the study that was most recently published was selected. After careful inspection, there were 18 studies involving overlapping samples. To ensure the independence of effect sizes, nine of these papers were included.

Coding of Studies

We coded the following information for each study included in the current meta-analysis: sample size, participants mean age, achievement level, types of achievement measurement, match between domain specificity of ASC and achievement, domain of measurement, time lags between assessments, variable types, sample types, publication year, and effect sizes.

Effect Sizes A cross-lagged regression coefficient generated by OSMASEM analysis that represented the effect of the predictor on the outcome variable after controlling for the initial levels of the outcome variable was used as effect size. Six Pearson correlation coefficients required for OSMASEM analysis needed to be coded. Among them were the concurrent correlation between two variables (i.e., the correlation of T1 ASC with T1 achievement, the correlation of T2 ASC with T2 achievement), the stability of ASC and achievement (i.e., the correlation of T1 ASC with T2 ASC, the correlation of T1 achievement with T2 achievement), and the cross-lagged correlation between ASC and achievement (i.e., the correlation of T1 ASC with T2 achievement, the correlation of T1 achievement with T2 ASC).

Age Since participant age differed at different time lags in the longitudinal studies included in the current meta-analysis, it was impossible to code their precise age. Instead, we coded the sample as childhood, adolescence, or adulthood according to different development phases. For studies that did not provide the students' age, we inferred this variable based on other valid information such as grade. For example, if a study adopted eighth-grade students as participants but did not report their age, we grouped them into adolescence.

Achievement Level Achievement level was coded as “typical-achieving” if the sample was unselected, “low-achieving” if the sample was described as at-risk or poor performance, or

“high-achieving” if the sample included students with high achievement, with excellent performance, or from the highest track.

Types of Achievement Measure We coded achievement as standardized tests, grades, or teacher-rating basing on how achievement was measured.

Match of Domain Specificity Between ASC and Achievement This moderator was coded as “match” if the measures of ASC and achievement belonged to the same subject (e.g., verbal self-concept and verbal achievement); otherwise, we coded it as “mismatch” (e.g., general ASC and mathematical achievement, mathematical self-concept, and GPA).

Domain of Measurement The domain was coded as “mathematical” if the relationship between mathematical self-concept and mathematics achievement was reported, or “verbal” if the relationship between self-concept and achievement in the verbal domain was reported.

Time Lags Between Measurement Waves We coded time lags between measurement waves in the following three categories: (a) less than or equal to 6 months; (b) more than 6 months and less than or equal to 12 months; and (c) more than 12 months. If a longitudinal study included more than two waves with an interval of several months, we coded it multiple times. For example, if a study included four waves with an interval of 6 months between each wave, it would be coded three times (T1–T2, time lag is 6 months; T1–T3, time lag is 12 months, and T1–T4, time lag is 18 months).

Variable Types This moderator was coded as “latent” if the reported effect sizes were based on the latent variable level, or “manifest” if the reported effect sizes were based on the manifest variable level.

Sample Types Sample types were coded as “representative sampling” if calculated effect sizes were based on a nationally representative sample, or “convenience sampling” if the calculated effect sizes were based on the sample selected by the researcher at a specific time and in a certain location.

Publication Year Since the latest publication year of the studies included in the Valentine et al.’s meta-analysis (2004) is 2001, we coded the studies with publication years before 2001 or equal to 2001 as “before 2001” and coded the studies after 2001 as “after 2001.”

Sample Sizes We used the sample sizes at T2 as the coding information. If some eligible studies only presented sample sizes at Time 1, we used other valid indicators such as the percent of attrition from the first assessment to the second assessment to estimate the sample sizes at Time 2.

In the coding process, sample sizes were coded as continuous variables that were centered for the analyses. Participant age, achievement level, types of achievement measurement, match or mismatch between domain specificity of ASC and achievement, domain of measurement, time lags between measurement waves, variable types, sample types and publication year were dummy-coded and submitted in *F* tests (as described below) to examine differences. See Online Resource 1 for summaries of the studies included in the meta-analysis.

Coder Reliability

All articles were coded by the first author, and 34 randomly selected articles were coded by the second author to provide estimates of interrater agreement. Interrater agreement was high ($ICC > .95$ and $k > .97$ for continuous and categorical variables, respectively) and discrepant ratings were discussed until consensus was reached.

Meta-analytic Procedure

The metaSEM package provides functions for univariate analysis, multivariate analysis, three-level meta-analysis, two-stage SEM (TSSEM), and OSMASEM using the SEM approach via the OpenMx package in the R statistical platform. The OSMASEM approach that we chose is the most suitable for processing longitudinal relationships between variables at continuous time points and building a cross-lagged model (Cheung, 2014). The metaSEM, with the maximum likelihood estimation for analyses, uses the sum rather than the average of sample sizes to compute the standard errors for the path coefficients, which increases the sensitivity of significance tests. For computation, we input the matrix of the sample size-weighted mean correlation in a dataset (Cheung & Chan, 2005).

To comprehensively test the direction of ASC and achievement, we separately analyzed the paths leading from ASC to achievement and from achievement to ASC. The effect size of each cross-lagged path was calculated based on the following equation (Cohen et al., 2013):

$$\beta_{Y1.2} = \frac{\gamma_{Y1} - \gamma_{Y2}\gamma_{12}}{1 - \gamma_{12}^2}$$

where $\beta_{Y1.2}$ is the standardized regression coefficient of X_1 predicting Y_2 after eliminating the effect of Y_1 on Y_2 . γ_{Y1} is the cross-lagged correlation between the predictor and the outcome (e.g., the correlation of T1 ASC with T2 achievement). γ_{Y2} is the autoregressive coefficient for the outcome variable (e.g., the relationship of T1 achievement with T2 achievement). And γ_{12} is the concurrent correlation between two variables, that is, the cross-sectional correlation (e.g., the correlation of T1 ASC with T1 achievement).

Referring to three previous meta-analyses (Khazanov & Ruscio, 2016; Kuykendall et al., 2015; Sowislo & Orth, 2013), cross-lagged effect sizes were also transformed using Fisher's Z_r transformations. The calculation and analysis of the effect sizes used the robust variance estimation (RVE) of the `robu()` function of the “`rubumeta`” package implemented in the R environment (version 3.3.3). RVE is a recently proposed meta-analytic method for dealing with dependent effect sizes (Fisher & Tipton, 2015). Because of the multidimensionality of ASC, most studies reported multiple correlated outcomes (e.g., the effect of mathematical achievement on mathematical self-concept; the effect of verbal achievement on verbal self-concept). Traditional meta-analytic methods are ill-equipped to deal with the complex and often unknown correlations among non-independent effect sizes, whereas RVE can provide valid point estimates, standard errors, and hypothesis tests even when the degree and structure of dependence between effect sizes is unknown (Fisher & Tipton, 2015; Hedges et al., 2010).

Given the high correlation between ASC and achievement, ρ was set to 0.80. ρ is the user-specified within-study effect size correlation, and its default value is 0.8 in RVE (Hedges et al., 2010). The user-specified value must be between 0 and 1 when using the correlation effect model (Fisher & Tipton, 2015). In fact, different settings of the ρ value would not significantly

influence final results. We also used the `Wald_test()` function in the `clubSandwich` package (version 0.2.2) in R to conduct F tests on models with dummy-coded coefficients, with the bias-reduced linearization adjustment for clustered standard errors and degrees of freedom estimated with Hotelling's T^2_Z method. In the effect size analyses, we used a random effects model that rests on the assumption that variation between studies is systematic and not only due to random error.

For each effect size, the z -score of the standardized residuals, which approximates a normal t distribution, was calculated to identify the presence of statistical outliers. If the z -score of the standardized residuals exceeded 1.96, the study was deemed to be an outlier (Viechtbauer & Cheung, 2010). Moreover, Cook's distance plots were inspected to examine the influence of outliers on the results. If the z -score of the standardized residuals exceeded 1.96 but the Cook's distance plots demonstrated that the outlier did not exert significant influence on the results, the study was retained.

To increase the robustness of the obtained effect size, we used three methods to test for publication bias. First, we visually inspected a funnel plot where the standard error for each study was plotted on the y -axis and the effect size on the x -axis. In the absence of publication bias, this plot would be expected to show a symmetrical shape. If the bottom of the funnel plot was asymmetrical, there would be evidence of publication bias. Second, we conducted Egger's regression intercept test for funnel plot asymmetry. Finally, we used the trim-and-fill procedure to calculate the possible number of missing studies based on asymmetry in the funnel plot and compute an estimated overall effect size on this basis.

Results

Description of Studies

The 68 studies included in the present meta-analysis were published between 1981 and 2020, with a median publication year in 2011. In total, 240 effect sizes were reported based on 72 independent samples. Sample sizes ranged from 43 to 14,985 ($M = 1,538$, $SD = 2,687$, *Median* [*Mdn*] = 546). The proportion of female participants ranged from 0 to 100%, with a mean of 47.17%. The average age of participants at the first assessment varied between 6.60 and 20.50 ($M = 12.19$, $SD = 2.92$, *Mdn* = 12.20). Time lags between assessments ranged from 3 to 66 months ($M = 20.62$, $SD = 13.46$, *Mdn* = 18). Most publications were conducted in the US (33.8%), Germany (17.6%), China (10.3%), Canada (5.9%), and Australia (5.9%). In total, 58.8% of studies explored the relationship between ASC and achievement in verbal or mathematical domains, while the remainder explored the relationship between general ASC and achievement.

Preliminary Analyses

Preliminary analyses identified two outliers in six analyses (for details, see Table 2). However, Cook's distance plots demonstrated that the outliers in all analyses did not have a significant influence on the results.

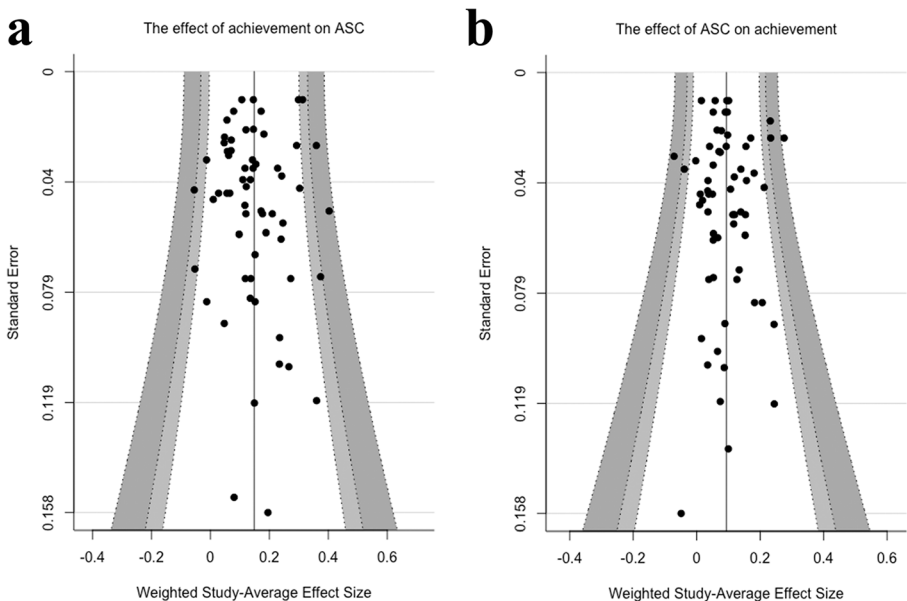
Since we were more interested in the REM of the relationship between ASC and achievement, we conducted publication bias tests on two cross-lagged paths. For the cross-lagged effect of T1

Table 2 Effect sizes for the relationship between ASC and achievement

Variable	<i>N</i>	<i>k</i>	Effect size	95% CI	<i>Q(df)</i>	<i>I</i> ²	<i>p</i>
T1 Concurrent correlation	68	72 ^a	0.40	[0.35, 0.44]	4873.78***(67)	97.99%	< 0.01
T2 Concurrent correlation	55	59	0.43	[0.38, 0.49]	2531.76***(54)	98.15%	< 0.01
Prospective, achievement → ASC	61	65	0.35	[0.31, 0.40]	2790***(60)	97.5%	< 0.01
Prospective, controlled achievement → ASC	61	65	0.15	[0.12, 0.18]	874.78***(60)	92.21%	< 0.01
Prospective, ASC → achievement	62	66 ^b	0.35	[0.31, 0.39]	2751.07***(61)	97.35%	< 0.01
Prospective, controlled ASC → achievement	62	66	0.09	[0.07, 0.11]	372.51***(61)	85.2%	< 0.01

Computations utilized a random effects model. *N* number of studies; *k* number of samples; *CI* confidence interval. ^a One outlier (Barker, 2007). ^b One outlier (Guo et al., 2015)

achievement on T2 ASC, an inspection of the funnel plot did not show asymmetry, indicating the absence of publication bias (see Fig. 2a). However, Egger's test for funnel plot asymmetry suggested possible publication bias, $t(59) = -2.53$, $p = 0.02$. The significant evidence for publication bias prompted a concern that if more null effects of achievement on ASC had been published, these studies might have reduced the effect to a trivial or negligible size. To address this concern, we used Duval and Tweedie's (2000) trim-and-fill procedure to look for missing studies on the left of the mean. However, the analysis did not estimate any missing studies (estimated missing = 0; $SE = 4.50$), showing that Egger's test may overestimate publication bias. For the cross-lagged effect of T1 ASC on T2 achievement, the funnel plot indicated that the distributions of effect sizes exhibited a symmetrical shape (see Fig. 2b). The Egger's test also returned nonsignificant results, $t(60) = -0.78$, $p = 0.44$. Similarly, trim-and-fill analysis showed no evidence of missing studies (estimated missing = 0; $SE = 4.30$).

**Fig. 2** a Funnel plot on the effect of achievement on ASC. b Funnel plot on the effect of ASC on achievement

Primary Analyses

First, we used the one-stage meta-analytic SEM to meta-analyze the studies in which ASC and achievement were all measured at least twice ($N = 55$). The results demonstrated that the concurrent correlations of ASC and achievement were 0.35, $p < 0.01$ at T1 and 0.13, $p < 0.01$ at T2. In addition, T1 ASC could significantly predict T2 ASC ($\beta = 0.45$, $p < 0.01$) and T1 achievement could significantly predict T2 achievement ($\beta = 0.64$, $p < 0.01$). Most importantly, ASC and achievement were reciprocally related, although the effect of achievement on ASC was stronger; that was, achievement predicted ASC, $\beta = 0.16$, $p < 0.01$; and ASC predicted achievement, $\beta = 0.08$, $p < 0.01$ (see Fig. 3), supporting the REM.

Additionally, to comprehensively test the REM, we separately analyzed the paths leading from achievement to ASC ($N = 61$) and from ASC to achievement ($N = 62$). The results supported the REM once more, and the effect of achievement on ASC ($\beta = 0.15$, $p < 0.01$, 95% CI [0.12, 0.18]) was larger than the effect of ASC on achievement ($\beta = 0.09$, $p < 0.01$, 95% CI [0.07, 0.11]) (see Table 2, Fig. 4a and b), strengthening our confidence in the reciprocal relationship between ASC and achievement. We examined the concurrent correlation at T1 and T2, and the prospective but uncontrolled correlation between two constructs, which allowed us to compare the meta-analytic results observed in previous research. As we hypothesized, when there was no control of the initial level of outcome, the concurrent correlation ($r_{T1} = 0.40$, [0.35, 0.44], $p < 0.01$; $r_{T2} = 0.43$, [0.38, 0.49], $p < 0.01$) and prospective correlation between ASC and achievement ($r_{\text{achievement} \rightarrow \text{ASC}} = 0.35$, [0.31, 0.40], $p < 0.01$; $r_{\text{ASC} \rightarrow \text{achievement}} = 0.35$, [0.31, 0.39], $p < 0.01$) increased to a far larger correlation. The weighted mean effect sizes are displayed in Table 2.

Moderator Analyses

In moderator analyses, we first focused the moderators of the effect of achievement on ASC. As expected, this effect was heterogeneous, $Q(60) = 874.78$, $p < 0.01$. Moreover, the percentage of between-study variance (I^2) was 92.21%, hinting that some variables may moderate the effect of achievement on ASC.

As expected, student age significantly moderated the effect of achievement on ASC, $t(54.5) = -2.26$, $p = 0.03$. The effect of achievement on ASC was stronger for children ($\beta = 0.18$, $p < 0.01$) than for adolescents ($\beta = 0.12$, $p < 0.01$), which indicated that the effect of achievement on ASC decreased as students grew older. Moderator analyses revealed no significant differences by student achievement level, $F(2, 1.83) = 1.27$, $p = 0.45$, types of achievement measurement, $F(2, 7.81) =$

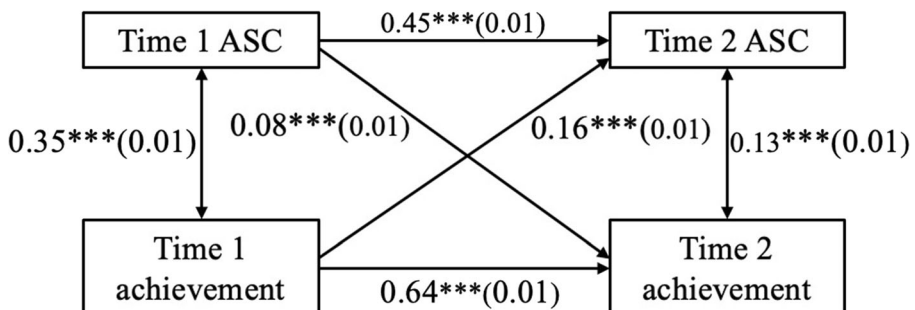


Fig. 3 Maximum likelihood estimation of the cross-lagged model

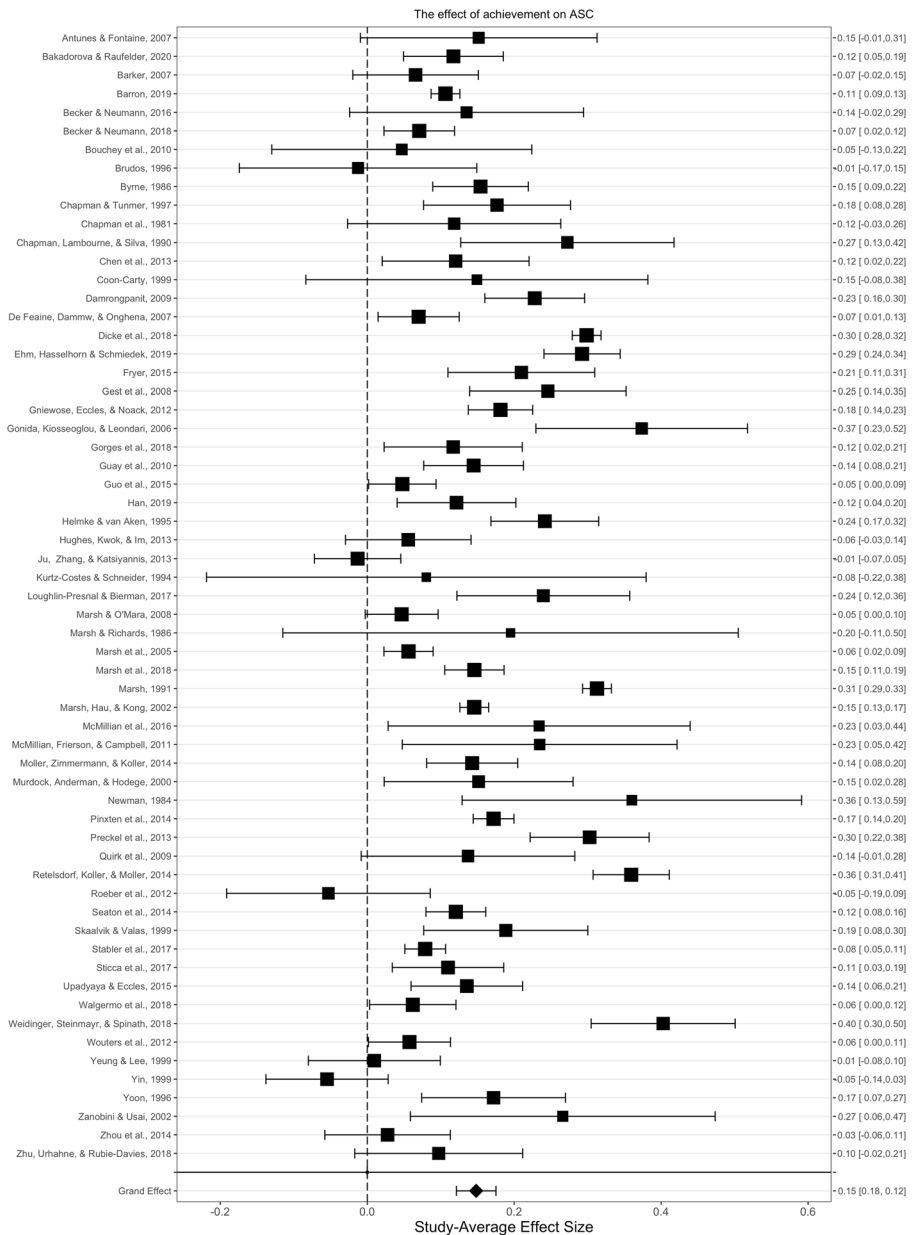


Fig. 4 a The forest plot of the effect of achievement on ASC. Size of the square indicates the relative weight assigned to that study in the analysis. Error bars represent the 95% confidence interval of the effect. This meta-analysis indicates that achievement could significantly predict the changes of ASC after controlling for the baseline level of ASC. b The forest plot of the effect of ASC on achievement. This meta-analysis indicates that ASC could significantly predict the changes of achievement after controlling for the baseline level of achievement

0.40, $p = 0.68$, match/mismatch of domain specificity between achievement and ASC, $t(56.7) = -0.90$, $p = 0.37$, or domain of measurement, $t(29) = -0.92$, $p = 0.36$.

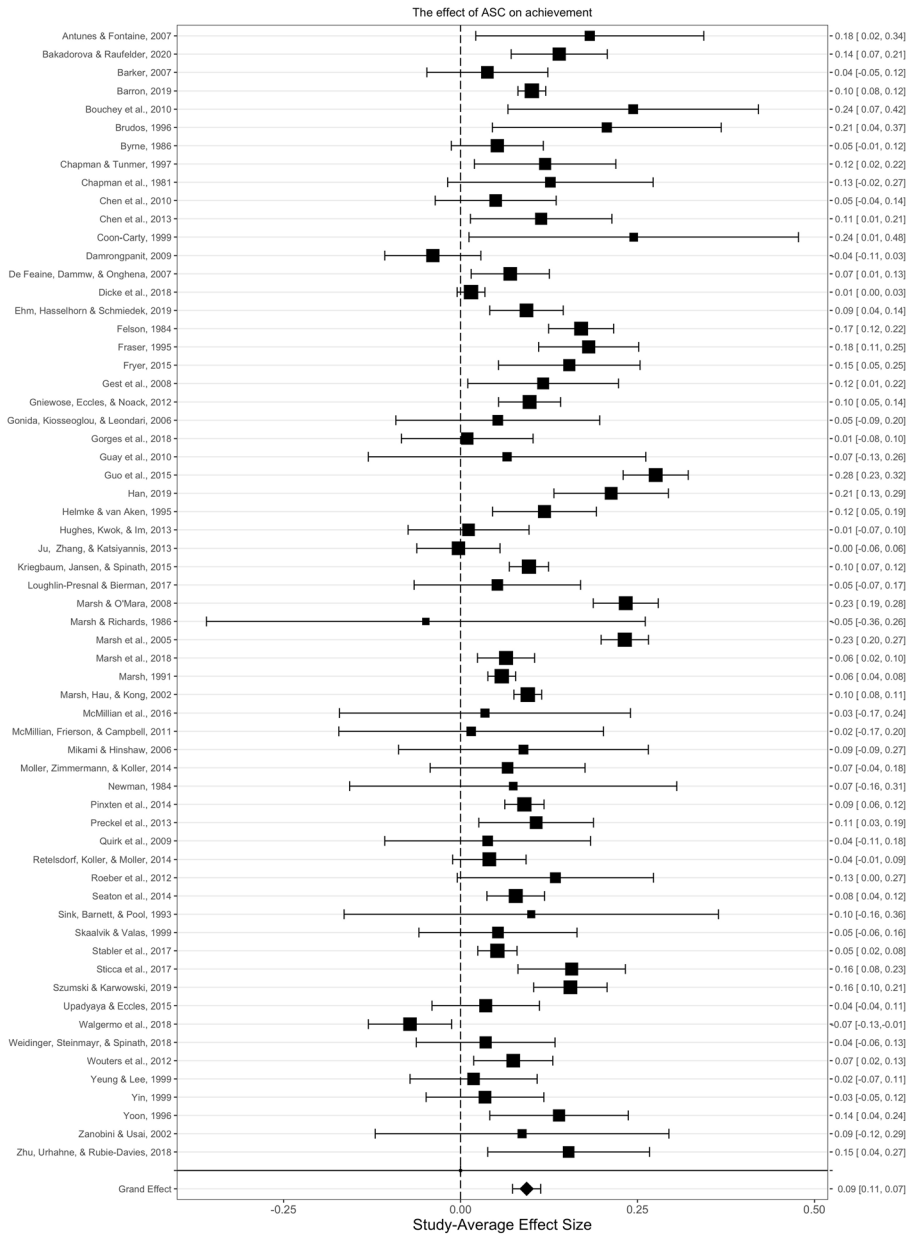


Fig. 4 (continued)

Finally, methodological variables did not significantly moderate the effect of achievement on ASC. Time lags between measurement waves, $F(2, 25.6) = 2.26, p = 0.13$, and variable types did not significantly moderate the effects of achievement on ASC, $t(40.5) = -0.23, p = 0.82$. No significant difference in effect sizes between representative and convenience

sampling were found, $t(38) = 0.14, p = 0.89$. The effect size in studies published before 2001 and after 2001 were equal, with no significant difference, $t(27) = 0.49, p = 0.63$ (see Table 3).

In addition, we conducted a moderator analysis of the path leading from ASC to achievement. The result of heterogeneity tests demonstrated the significant heterogeneity between effect sizes ($Q(61) = 372.51, p < 0.01$), and the percentage of between-study variance (I^2) was 85.20%.

Three variables emerged as significant moderators of the effect of ASC on achievement. The first was student age, $t(49) = 2.85, p = 0.006$. The effect of ASC on achievement was stronger for adolescents ($\beta = 0.11, p < 0.01$) than for children ($\beta = 0.06, p < 0.01$), which was contrary to the effect of achievement on ASC. The second significant moderator was student achievement level, $t(10.2) = -4.51, p = 0.01$. The effect sizes for low-achieving students ($\beta = 0.01, p < 0.01$) were significantly weaker than typical-achieving students ($\beta = 0.11, p < 0.01$). Moreover, types of achievement measurement significantly moderated the effect of ASC on achievement, $F(2, 4.88) = 6.74, p = 0.04$. In line with our hypothesis, when achievement measurement used standardized tests, the effect of ASC on achievement was minimal ($\beta = 0.06, p < 0.01$), compared to grades ($\beta = 0.13, p < 0.01$) and teacher-rating ($\beta = 0.10, p < 0.01$).

Table 3 Moderations on the effect of achievement on ASC

Moderator	<i>t/F</i>	Effect size	<i>df</i>	<i>p</i>
Age	-2.26		54.5	0.03
Childhood		0.18	27.7	< 0.01
Adolescence		0.12	28.3	< 0.01
Adulthood		NA		
Achievement level	1.27		1.83	0.45
Low-achieving		0.10	6.42	< 0.01
Typical-achieving		0.15	47.14	< 0.01
High-achieving		0.23	1	< 0.01
Types of achievements measurement	0.40		7.81	0.68
Standard tests		0.14	26.43	< 0.01
Grades		0.16	31.84	< 0.01
Teacher-rating		0.18	3	0.04
Match between domain specificity	-0.90		56.7	0.37
Match		0.16	28.7	< 0.01
Mismatch		0.14	28.9	< 0.01
Domain	-0.92		29	0.36
Mathematics		0.17	16.4	< 0.01
Language		0.14	17.1	< 0.01
Time lags	2.26		25.6	0.13
≤ 6 months		0.09	9.05	0.02
> 6 months ≤ 12 months		0.14	28.1	< 0.01
> 12 months		0.17	33.5	< 0.01
Variable types	-0.23		40.5	0.82
Latent variable		0.14	37.6	< 0.01
Manifest variable		0.15	18.8	< 0.01
Sample types	0.14		38	0.89
Representative sampling		0.15	37	< 0.01
Convenience sampling		0.15	19.6	< 0.01
Publication year	0.49		27	0.63
After 2001		0.14	41.1	< 0.01
Before 2001		0.16	15.5	< 0.01

Significant ($p < 0.10$) moderators are shown in boldface font. If $df < 4$, there is a twofold greater risk of making a Type I error. The listed p value represents the significance of the moderator in question. When there were not enough studies to estimate an effect, NA is listed in the column

We then found that match or mismatch of domain specificity between achievement and ASC ($t(53.7) = 0.94, p = 0.35$), and domain of measurement, ($t(27.5) = -0.85, p = 0.40$) did not moderate the effect of ASC on achievement. In addition, no methodological variables had a significant moderating effect on effect sizes. Time lags between measurement waves did not moderate the effect of ASC on achievement, $F(2, 21.6) = 0.98, p = 0.39$. The use of latent variables did not moderate the effect of ASC on achievement, $t(35.1) = 1.04, p = 0.30$. The effect sizes did not vary according to sample types ($t(31.8) = -0.68, p = 0.50$) or publication year ($t(23.6) = 0.68, p = 0.50$) (see Table 4).

For comparison with Valentine et al.'s meta-analysis (2004), we included studies that reported regression coefficients to conduct primary effects and moderator analyses. A more detailed analysis is presented in the *Online Resource 2*.

Difference in the Relationship Between ASC and Achievement

Primary analyses and moderator analyses found that the effect of ASC on achievement and the effect of achievement on ASC were significant both in childhood and adolescence, supporting the generalizability of the REM across age. However, combining the moderator analyses of the effect of ASC on achievement and achievement on ASC, it was found that when students were

Table 4 Moderations on the effect of ASC on achievement

Moderator	<i>t/F</i>	Effect size	<i>df</i>	<i>p</i>
Age	2.85		49	0.006
Childhood		0.06	23.9	< 0.01
Adolescence		0.11	29.7	< 0.01
Adulthood		NA		
Achievement level	-4.51		10.2	0.01
Low-achieving		0.01	7.55	< 0.01
Typical-achieving		0.11	44.1	< 0.01
High-achieving		NA		
Types of achievement measurement	6.74		4.88	0.04
Standard tests		0.06	30.12	< 0.01
Grades		0.13	25.13	< 0.01
Teacher-rating		0.10	1.99	0.01
Match between domain specificity	0.94		53.7	0.35
Match		0.08	27.3	< 0.01
Mismatch		0.10	27.2	< 0.01
Domain	-0.85		27.5	0.40
Mathematics		0.09	15.9	< 0.01
Language		0.07	15.5	< 0.01
Time lags	0.98		21.6	0.39
≤ 6 months		0.07	7.49	0.11
> 6 months ≤ 12 months		0.08	28	< 0.01
> 12 months		0.11	32.6	< 0.01
Variable types	1.04		35.1	0.30
Latent variable		0.11	16.6	< 0.01
Manifest variable		0.08	37.1	< 0.01
Sample types	-0.68		31.8	0.50
Representative sampling		0.09	36.6	< 0.01
Convenience sampling		0.10	17.8	< 0.01
Publication year	0.68		23.6	0.50
After 2001		0.09	40.1	< 0.01
Before 2001		0.10	13.9	< 0.01

in childhood, the effect of achievement on ASC ($\beta = 0.18, p < 0.01, 95\% \text{ CI } [0.15, 0.22]$) was stronger than the effect of ASC on achievement ($\beta = 0.06, p < 0.01, 95\% \text{ CI } [0.04, 0.08]$); when students entered adolescence, the effect of achievement on ASC ($\beta = 0.12, p < 0.01, 95\% \text{ CI } [0.09, 0.16]$) was almost the same as it was vice-versa ($\beta = 0.11, p < 0.01, 95\% \text{ CI } [0.08, 0.14]$) (see Fig. 5), demonstrating a trend from a strong skill-development effect to a pronounced reciprocal effect with age within the framework of the REM.

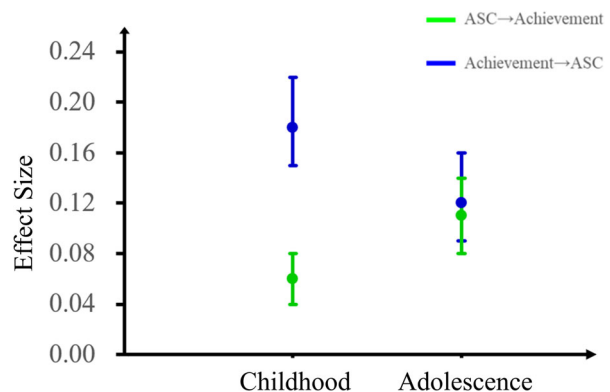
Discussion

The current meta-analysis explored the longitudinal relationship between ASC and achievement and its moderators by a focus upon available studies. Consistent with previous longitudinal studies (e.g., Wigfield & Eccles, 2000; Marsh et al., 2002; Gorges et al., 2018; Weidinger et al., 2018; Sewasew & Schroeders, 2019), our results found that ASC and achievement affected each other mutually within domains over time, supporting the REM.

However, the mutual dependency did not indicate that effect sizes were equal: the skill-development effect was stronger than the self-enhancement effect. The small effect sizes of the REM may be related to the moderate stability of ASC ($r = 0.45$) and achievement ($r = 0.64$) (Seaton et al., 2015). The REM suggests that students are likely to feel more competent in subjects in which they perform better, and the sense of competence also affects subsequent achievement. As Prast et al. (2018) found, perceived competence could significantly predict T3 achievement and played a partial mediating role in the association between previous and subsequent achievement. Our findings corroborated the argument that a positive ASC is an important educational outcome and a mediating variable that promotes the development of achievement (Marsh & Martin, 2011; Möller et al., 2009).

Moderator analyses revealed that the effect of ASC on achievement and the effect of achievement on ASC were moderated by student age, although the support for the REM was found in both childhood and adolescence. The nature of the relationship between ASC and achievement demonstrated a trend from a strong skill-development effect to a more pronounced reciprocal effect with age within the framework of the REM, which was inconsistent with the results of Valentine et al.'s meta-analysis (2004). However, it should be noted that in Valentine et al.'s meta-analysis (2004), more than half of the included studies used middle school students as samples, perhaps limiting the comparison of effect sizes at different developmental stages.

Fig. 5 The development of the relationship between ASC and achievement from childhood to adolescence



What factors might explain this developmental change? During elementary school, students' ASC undergoes a process of shaping and reshaping that is dominated by cumulative achievement-related success and failure (Skaalvik & Hagtvet, 1990). The motivational properties of ASC have not been fully developed; thus, their influence on achievement is limited (Helmke & van Aken, 1995). However, ASC is expected to exert increasing effects on achievement or other academic choices when it is firmly established and more stable (Wigfield & Karpathian, 1991). For instance, Guay et al. (2003) have found that the reliability and stability of ASC improved with students' age, which increased the potential for influencing subsequent achievement. The developmental change of the relationship between ASC and achievement also supports the hypothesis that self-beliefs and achievement gradually become a coherent self-stabilizing system with age (Dweck, 2002).

Age differences in the relationship between ASC and achievement might be related to the maturation of cognitive ability, which enables students to make inferences regarding their own and others' competence by integrating multiple information resources such as performance feedback and reflected appraisals from significant others (Weidinger et al., 2018). In addition, high cognitive abilities could enhance the coordination between previous self-beliefs, thus making it more possible for self-concept to predict external achievement indicators (Skaalvik & Hagtvet, 1990). Furthermore, significant changes have occurred in the educational environment for students beyond elementary school, with students having to participate in an increasing number of normative evaluations and receive more judgments from teachers (Dweck, 2002; Stipek & Iver, 1989). Such feedback helps students to establish competence beliefs more firmly, thereby boosting the effect of ASC on achievement. Taken together, such modes of progress lead to a strong skill-development effect for younger children but a pronounced reciprocal effect for older students.

In addition, we found that achievement level significantly moderated the effect of ASC on achievement. The effect of ASC on achievement for low-achieving students was significantly weaker than typical-achieving students. Low-achieving students experience less success and more academic challenges, which can lead to their academic motivation being weakened and the effect of ASC on achievement being greatly reduced (Hampton & Mason, 2003; Prast et al., 2018). Nevertheless, we cannot ignore the implications of ASC for achievement for low-achieving students. Instead, high ASC might be especially beneficial for subsequent achievement among low-achieving students because the motivational properties of ASC could initiate adaptive learning strategies and behaviors such as efforts and persistence, which would have a positive effect on future achievement and success (Marsh, 1990; Marsh & Martin, 2011; Wigfield & Eccles, 2000).

Consistent with our hypothesis, type of achievement measurement was a significant moderator of the effect of ASC on achievement. Studies using standardized tests as indicators of achievement tended to report a weaker relationship between ASC and achievement than studies using grades or teacher-rating as indicators. As mentioned in the introduction, standardized testing is not high-stakes and not of personal significance for the students, while grades and teacher-rating often serve other functions such as motivating students (Kriegbaum et al., 2018; Pinxten et al., 2010). ASC possesses motivational properties, so this leads to the closer correlation between ASC and grade or teacher-rating as they are directly related to motivation (Valentine et al., 2004; Wylie, 1974).

Moderator analyses demonstrated that the match between domain specificity of ASC and achievement did not significantly affect the relationship between ASC and achievement, which was inconsistent with Valentine et al.'s meta-analysis (2004). In the present study, we coded

this moderator as mismatch when general self-concept corresponded to mathematical/verbal achievement, or the mathematical/verbal self-concept corresponded to GPA. In fact, the general ASC and mathematics achievement are not totally mismatched. However, in Valentine et al.'s meta-analysis (2004), the researchers coded the moderator as mismatch if ASC and achievement were based on different subjects (e.g., mathematical self-concept corresponded to verbal achievement). Therefore, it was not surprising that the significant moderating effect of match between domain specificity of ASC and achievement was not found in our study.

Similarly, the moderating effect of domain was not found in the present study. Although Chen et al. (2013) and Möller et al. (2014) found that a reciprocal relationship for mathematical but not for verbal domain, some characteristics of two studies needed to be noted. For instance, the stability of verbal achievement ($r = .75$) was much higher in Chen et al.'s research (2013), and the time lags between T1 and T2 was as long as 4 years in Moller et al.'s (2014) study, which may have resulted in the masking of the effect of the prior verbal self-concept. Language and reading are often viewed as feminine domains, whereas mathematics is stereotyped as a masculine domain (Marsh, 1989). Skaalvik and Skaalvik (2014) revealed that male students had a significantly higher ASC, performance expectation, and intrinsic motivation for mathematics than female students. The differential socialization hypothesis also states that gender-related difference in socialization progress may not fully enhance girls' self-concept and performance in the mathematical domain and may fail to reinforce boys' self-concept and performance in the verbal domain (Marsh, 1989). Therefore, future research is needed to investigate the interaction of subject and gender on the relationship between ASC and achievement.

Furthermore, we found that the effect of achievement on ASC and the effect of ASC on achievement did not significantly vary as a function of time lags. Our results suggested that the skill-development effect and self-enhancement effect are detectable and stable across a wide range of time lags. Nevertheless, although our meta-analysis covered a wide range of time lags (from 2 to 66 months), the number of studies with time lags of less than 1 year was too small, which may restrict the statistical power. Thus, these results are merely tentative and should not be regarded as definitive evidence. Future research should further investigate the effect of time lags on the relationship between ASC and achievement, and the mediating mechanisms that account for the temporal stability of these effects. Although there is no general standard for optimal measurement intervals, Helmke and van Aken (1995) suggested that students' age and school system should be considered when determining the time interval between measurement waves.

Although some studies maintained that calculation effect sizes at the latent level within SEM can yield more valid and potentially high estimates, the use of latent variables did not cause the relationship between ASC and achievement to be significantly higher than those using manifest variables (Leonhart et al., 2008; Valentine et al., 2004). As for sample type of the original study, although representative sampling has some advantages over convenience sampling, their limitations are clear. In general, representative sampling has relatively broader aims, so that some variable of interest cannot be incorporated. For example, Project for the Analysis of Learning and Achievement in Mathematics only targeted students' mathematical domain, and the Youth in Transition Project only measured students' general ASC, which pertained specifically to a single academic domain. Therefore, it was not surprising that no significant difference in effect sizes was found between representative and convenience sampling. In the same way, longitudinal studies after 2001 also supported the reciprocal relationship between of ASC and achievement, strengthening the robustness of REM.

Limitations and Future Directions

Although the present meta-analysis of longitudinal studies has gone a step beyond previous studies, several limitations should be borne in mind when interpreting certain results. First, our research did not allow for a causal inference regarding the relationship between ASC and achievement because the included studies all adopted a correlational design. Therefore, we cannot rule out a third variable that affects both constructs, particularly due to the modest effect sizes observed here. For example, academic interest is related to high achievement and ASC (Marsh et al., 2005), but the extent to which academic interest accounts for their relationship is unclear. Future investigation could use experimental design, quasi-experimental design, and invention strategies to directly test the causal ordering between achievement and ASC.

Second, although the present study found that the relationship between achievement and ASC varied as a function of age and presented some evidence for the changes of the relationship between ASC and achievement over time, it did not provide a full picture of the developmental trajectory of the relationship between two constructs. Therefore, future studies could follow the same cohort of participants over time from childhood to adolescence to derive a linear or non-linear trajectory of the relationship between ASC and achievement from childhood to adolescence via longitudinal studies with a multicohort–multioccasion design. In addition, although we have inferred that this development may owe both to improvement of students' cognitive abilities and changes of educational contexts, more in-depth research is needed to explore the mechanism underlying the change of relationship between ASC and achievement.

Third, our study is limited in that we did not systematically search for unpublished studies. Although we searched for unpublished studies to reduce publication bias, we did not request the data on unpublished studies from researchers who had published many studies in the field, which might have resulted in some unpublished data not being included.

Finally, abundant research has revealed that ASC and academic self-efficacy conceptually and empirically represent two distinct constructs (e.g., Bong & Skaalvik, 2003; Ferla et al., 2009; Pietsch et al., 2003), so we only included longitudinal studies where ASC or achievement was measured at least twice. However, it is undeniable that we failed to contrast the difference between ASC and academic self-efficacy when focusing only on ASC.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10648-021-09600-1>.

Acknowledgements The authors would like to acknowledge Herbert W. Marsh and Fang Xie for their helpful comments in relation to this article.

Author Contribution Conceptualization: Huimin Wu, Cheng Guo and Yiqun Guo; methodology: Huimin Wu, Yingkai Yang and Yiqun Guo; writing—original draft preparation: Huimin Wu and Yiqun Guo; writing—review and editing: Le Zhao and Cheng Guo; funding acquisition: Yiqun Guo, Le Zhao, and Cheng Guo.

Declarations

Conflict of Interest The authors declare that they have no conflict of interest.

References

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

- Acharya, A. S., Prakash, A., Saxena, P., & Nigam, A. (2013). Sampling: Why and how of it. *Indian Journal of Medical Specialties*, 4(2), 330–333. <https://doi.org/10.7713/ijms.2013.00331>.
- Albert, S. (1977). Temporal comparison theory. *Psychological Review*, 84(6), 485–503. <https://doi.org/10.1037/0033-295X.84.6.485>.
- *Antunes, C., & Fontaine, A. (2007). Gender differences in the causal relationship between adolescents' mathematics self-concept and scholastic performance. *Psychologica Belgica*, 47(1), 71–94. <https://doi.org/10.5334/pb-47-1-71>.
- Arens, A. K., Bodkin-Andrews, G., Craven, R. G., & Yeung, A. S. (2014). Self-concept of indigenous and non-indigenous Australian students: Competence and affect components and relations to achievement. *Learning and Individual Differences*, 32, 93–103. <https://doi.org/10.1016/j.lindif.2014.03.019>.
- Arens, A. K., & Hasselhorn, M. (2015). Differentiation of competence and affect self-perceptions in elementary school students: Extending empirical evidence. *European Journal of Psychology of Education*, 30(4), 405–419. <https://doi.org/10.1007/s10212-015-0247-8>.
- *Bakadorova, O., & Raufelder, D. (2020). The relationship of school self-concept, goal orientations and achievement during adolescence. *Self and Identity*, 19(2), 235–249. <https://doi.org/10.1080/15298868.2019.1581082>.
- Bandura, A. (2010). Self-efficacy. *The Corsini Encyclopedia of Psychology*, 1–3.
- *Barker, K. L. (2007). Students' goals, academic self-concept and academic achievement: Testing competing models of causation. In *Research Impacts: Proving or Improving? Proceedings of the AARE International Education Research Conference, 25-29 November 2007, Fremantle, WA*.
- *Barron, C. M. (2019). *Longitudinal investigation into the role of reading self-concept on reading achievement for English as a first language and English language learners*. University of Toronto.
- *Becker, M., & Neumann, M. (2016). Context-related changes in academic self-concept development: On the long-term persistence of big-fish-little-pond effects. *Learning and Instruction*, 45, 31–39. <https://doi.org/10.1016/j.learninstruc.2016.06.003>.
- *Becker, M., & Neumann, M. (2018). Longitudinal big-fish-little-pond effects on academic self-concept development during the transition from elementary to secondary schooling. *Journal of Educational Psychology*, 110(6), 882–897. <https://doi.org/10.1037/edu0000233>.
- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, 15(1), 1–40. <https://doi.org/10.1023/A:1021302408382>.
- *Bouchey, H. A., Shoulberg, E. K., Jodl, K. M., & Eccles, J. S. (2010). Longitudinal links between older sibling features and younger siblings' academic adjustment during early adolescence. *Journal of Educational Psychology*, 102(1), 197–211. <https://doi.org/10.1037/a0017487>.
- *Brudos, G. K. (1996). *The relationships among behavior problems, academic achievement, and achievement-related beliefs in middle school-aged children: a longitudinal study*. The Florida State University.
- *Byrne, B. M., & Shavelson, R. J. (1986). On the structure of adolescent self-concept. *Journal of Educational Psychology*, 78(6), 474–481. <https://doi.org/10.1037/0022-0663.78.6.474>.
- Calsyn, R. J., & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? *Journal of Educational Psychology*, 69(2), 136–145. <https://doi.org/10.1037/0022-0663.69.2.136>.
- *Chapman, J. W., Cullen, J. L., Boersma, F. J., & Maguire, T. O. (1981). Affective variables and school achievement: A study of possible causal influences. *Canadian Journal of Behavioural Science*, 13(2), 181–192. <https://doi.org/10.1037/h0081169>.
- *Chapman, J. W., Lambourne, R., & Silva, P. A. (1990). Some antecedents of academic self-concept: A longitudinal study. *British Journal of Educational Psychology*, 60(2), 142–152. <https://doi.org/10.1111/j.2044-8279.1990.tb00931.x>.
- *Chapman, J. W., & Tunmer, W. E. (1997). A longitudinal study of beginning reading achievement and reading self-concept. *British Journal of Educational Psychology*, 67(3), 279–291. <https://doi.org/10.1111/j.2044-8279.1997.tb01244.x>.
- *Chen, Q., Hughes, J. N., Liew, J., & Kwok, O. M. (2010). Joint contributions of peer acceptance and peer academic reputation to achievement in academically at-risk children: Mediating processes. *Journal of Applied Developmental Psychology*, 31(6), 448–459. <https://doi.org/10.1016/j.appdev.2010.09.001>.
- *Chen, S.-K., Yeh, Y.-C., Hwang, F.-M., & Lin, S. S. (2013). The relationship between academic self-concept and achievement: A multicohort-multioccasion study. *Learning and Individual Differences*, 23, 172–178. <https://doi.org/10.1016/j.lindif.2012.07.021>.

- Cheung, M. W.-L. (2014). Fixed-and random-effects meta-analytic structural equation modeling: Examples and analyses in R. *Behavior Research Methods*, *46*(1), 29–40. <https://doi.org/10.3758/s13428-013-0361-y>.
- Cheung, M. W.-L., & Chan, W. (2005). Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*, *10*(1), 40–64. <https://doi.org/10.1037/1082-989X.10.1.40>.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.
- Cole, D. A., & Maxwell, S. E. (2003). Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. *Journal of Abnormal Psychology*, *112*(4), 558–577. <https://doi.org/10.1037/0021-843X.112.4.558>.
- Collins, L. M., & Graham, J. W. (2002). The effect of the timing and spacing of observations in longitudinal studies of tobacco and other drug use: Temporal design considerations. *Drug and Alcohol Dependence*, *68*, 85–96. [https://doi.org/10.1016/S0376-8716\(02\)00217-X](https://doi.org/10.1016/S0376-8716(02)00217-X).
- *Coon-Carty, H. M. (1999). *The influence of subject-specific academic self-concepts on academic achievement from a Vygotskian developmental perspective: A longitudinal investigation*. Loyola University of Chicago Adviser.
- *Damrongpanit, S. (2009). The study of growth between academic self-concept, nonacademic self-concept, and academic achievement of ninth-grade students: A multiple group analysis. *Research in Higher Education Journal*, *5*, 1–13.
- *De Fraine, B., Van Damme, J., & Onghena, P. (2007). A longitudinal analysis of gender differences in academic self-concept and language achievement: A multivariate multilevel latent growth approach. *Contemporary Educational Psychology*, *32*(1), 132–150. <https://doi.org/10.1016/j.cedpsych.2006.10.005>.
- Denissen, J. J., Zaretz, N. R., & Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: Longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Development*, *78*(2), 430–447. <https://doi.org/10.1111/j.1467-8624.2007.01007.x>.
- *Dicke, T., Marsh, H. W., Parker, P. D., Pekrun, R., Guo, J., & Televantou, I. (2018). Effects of school-average achievement on individual self-concept and achievement: Unmasking phantom effects masquerading as true compositional effects. *Journal of Educational Psychology*, *110*(8), 1112–1126. <https://doi.org/10.1037/edu0000259>.
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. <https://doi.org/10.1111/j.0006-341X.2000.00455.x>.
- Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist*, *41*, 1040–1048. <https://doi.org/10.1037/0003-066X.41.10.1040>.
- Dweck, C. S. (2002). The development of ability conceptions. In A. Wigfield & J. Eccles (Eds.), *The development of achievement motivation*. Academic Press.
- *Ehm, J. H., Hasselhorn, M., & Schmiedek, F. (2019). Analyzing the developmental relation of academic self-concept and achievement in elementary school children: Alternative models point to different results. *Developmental Psychology*, *55*(11), 2336–2351. <https://doi.org/10.1037/dev0000796>.
- Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1–4. <https://doi.org/10.11648/j.ajtas.20160501.11>.
- Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences*, *19*(4), 499–505. <https://doi.org/10.1016/j.lindif.2009.05.004>.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140. <https://doi.org/10.1177/001872675400700202>.
- *Felson, R. B. (1984). The effect of self-appraisals of ability on academic performance. *Journal of Personality and Social Psychology*, *47*(5), 944–952. <https://doi.org/10.1037/0022-3514.47.5.944>.
- Fisher, Z., & Tipton, E. (2015). Robumeta: An R-package for robust variance estimation in meta-analysis. Retrieved from <http://arxiv.org/abs/1503.02220>.
- *Fraser, K. L. (1995). *Hopes, fears, and the architecture of adolescence: the possible selves as polaroids and blueprints of early adolescent development*. University of Colorado.
- *Fryer, L. K. (2015). Predicting self-concept, interest and achievement for first-year students: The seeds of lifelong learning. *Learning and Individual Differences*, *38*, 107–114. <https://doi.org/10.1016/j.lindif.2015.01.007>.
- *Gest, S. D., Rulison, K. L., Davidson, A. J., & Welsh, J. A. (2008). A reputation for success (or failure): The association of peer academic reputations with academic self-concept, effort, and performance across the upper elementary grades. *Developmental Psychology*, *44*(3), 625–636. <https://doi.org/10.1037/0012-1649.44.3.625>.

- *Gniewosz, B., Eccles, J. S., & Noack, P. (2012). Secondary school transition and the use of different sources of information for the construction of the academic self-concept. *Social Development, 21*(3), 537–557. <https://doi.org/10.1111/j.1467-9507.2011.00635.x>.
- *Gonida, E., Kiosseoglou, G., & Leondari, A. (2006). Implicit theories of intelligence, perceived academic competence, and school achievement: Testing alternative models. *The American Journal of Psychology, 119*(2), 223–238. <https://doi.org/10.2307/20445336>.
- *Gorges, J., Neumann, P., Wild, E., Stranghöner, D., & Lütje-Klose, B. (2018). Reciprocal effects between self-concept of ability and performance: A longitudinal study of children with learning disabilities in inclusive versus exclusive elementary education. *Learning and Individual Differences, 61*, 11–20. <https://doi.org/10.1016/j.lindif.2017.11.005>.
- Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., Oliver, P. H., & Guerin, D. W. (2007). Multivariate latent change modeling of developmental decline in academic intrinsic mathematics motivation and achievement: Childhood through adolescence. *International Journal of Behavioral Development, 31*, 317–327. <https://doi.org/10.1177/0165025407077752>.
- Götz, T., Cronjäger, H., Frenzel, A. C., Lüdtke, O., & Hall, N. C. (2010). Academic self-concept and emotion relations: Domain specificity and age effects. *Contemporary Educational Psychology, 35*(1), 44–58. <https://doi.org/10.1016/j.cedpsych.2009.10.001>.
- Graham, S., & Weiner, B. (1996). Theories and principles of motivation. *Handbook of Educational Psychology, 4*, 63–84.
- Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology, 95*, 124–136. <https://doi.org/10.1037/0022-0663.95.1.124>.
- *Guay, F., Ratelle, C. F., Roy, A., & Litalien, D. (2010). Academic self-concept, autonomous academic motivation, and academic achievement: Mediating and additive effects. *Learning and Individual Differences, 20*(6), 644–653. <https://doi.org/10.1016/j.lindif.2010.08.001>.
- *Guo, J., Marsh, H. W., Morin, A. J., Parker, P. D., & Kaur, G. (2015). Directionality of the associations of high school expectancy-value, aspirations, and attainment: A longitudinal study. *American Educational Research Journal, 52*(2), 371–402. <https://doi.org/10.3102/0002831214565786>.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). *Multivariate data analysis (Vol 5, No. 3, pp. 207-219)*. Prentice hall.
- Hampton, N. Z., & Mason, E. (2003). Learning disabilities, gender, sources of efficacy, self-efficacy beliefs, and academic achievement in high school students. *Journal of School Psychology, 41*(2), 101–112. [https://doi.org/10.1016/S0022-4405\(03\)00028-1](https://doi.org/10.1016/S0022-4405(03)00028-1).
- *Han, F. (2019). Longitudinal relations between school self-concept and academic achievement. *Revista de Psicodidáctica (English ed.)*, *24*(2), 95–102. <https://doi.org/10.1016/j.psicoe.2019.03.002>.
- Harter, S. (1978). Effectance motivation reconsidered. Toward a developmental model. *Human Development, 21*(1), 34–64. <https://doi.org/10.1159/000271574>.
- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*(1), 39–65. <https://doi.org/10.1002/jrsm.17>.
- *Helmke, A., & van Aken, M. A. (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>.
- Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relationships. *Journal of School Psychology, 49*(5), 505–528. <https://doi.org/10.1016/j.jsp.2011.07.001>.
- *Hughes, J. N., Kowk, O. M., & Im, M. (2013). Effect of retention in first grade on parents' educational expectations and children's academic outcomes. *American Educational Research Journal, 50*(6), 1336–1359. <https://doi.org/10.3102/0002831213490784>.
- *Ju, S., Zhang, D., & Katsiyannis, A. (2013). The causal relationship between academic self-concept and academic achievement for students with disabilities: An analysis of SEELS data. *Journal of Disability Policy Studies, 24*(1), 4–14. <https://doi.org/10.1177/1044207311427727>.
- Kadir, M. S., Yeung, A. S., & Diallo, T. M. (2017). Simultaneous testing of four decades of academic self-concept models. *Contemporary Educational Psychology, 51*, 429–446. <https://doi.org/10.1016/j.cedpsych.2017.09.008>.
- Khazanov, G. K., & Ruscio, A. M. (2016). Is low positive emotionality a specific risk factor for depression? A meta-analysis of longitudinal studies. *Psychological Bulletin, 142*(9), 991–1015. <https://doi.org/10.1037/bul0000059>.
- Kriegbaum, K., Becker, N., & Spinath, B. (2018). The relative importance of intelligence and motivation as predictors of school achievement: A meta-analysis. *Educational Research Review, 25*, 120–148. <https://doi.org/10.1016/j.edurev.2018.10.001>.

- *Kriegbaum, K., Jansen, M., & Spinath, B. (2015). Motivation: A predictor of PISA's mathematics competence beyond intelligence and prior test achievement. *Learning and Individual Differences, 43*, 140–148. <https://doi.org/10.1016/j.lindif.2015.08.026>.
- *Kurtz-Costes, B. E., & Schneider, W. (1994). Self-concept, attributional beliefs, and school achievement: A longitudinal analysis. *Contemporary Educational Psychology, 19*(2), 199–216. <https://doi.org/10.1006/ceps.1994.1017>.
- Kuykendall, L., Tay, L., & Ng, V. (2015). Leisure engagement and subjective well-being: A meta-analysis. *Psychological Bulletin, 141*(2), 364–403. <https://doi.org/10.1037/a0038508>.
- Leonhart, R., Wirtz, M., & Bengel, J. (2008). Measuring effect sizes using manifest versus latent variables: Consequences and implications for research. *International Journal of Rehabilitation Research, 31*(3), 207–216. <https://doi.org/10.1097/MRR.0b013e3282fb7872>.
- Liem, G. A. D., McInerney, D. M., & Yeung, A. S. (2015). Academic self-concepts in ability streams: Considering domain specificity and same-stream peers. *The Journal of Experimental Education, 83*(1), 83–109. <https://doi.org/10.1080/00220973.2013.876227>.
- Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path, and structural equation analysis* (4th ed.). New Jersey: Guilford.
- *Loughlin-Prernal, J., & Bierman, K. L. (2017). How do parent expectations promote child academic achievement in early elementary school? A test of three mediators. *Developmental Psychology, 53*(9), 1694–1708. <https://doi.org/10.1037/dev0000369>.
- Marsh, H. W. (1989). Age and sex effects in multiple dimensions of self-concept: Preadolescence to early adulthood. *Journal of Educational Psychology, 81*(3), 417–430. <https://doi.org/10.1037/0022-0663.81.3.417>.
- Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*(4), 646–656. <https://doi.org/10.1037/0022-0663.82.4.646>.
- *Marsh, H. W. (1991). Failure of high-ability high schools to deliver academic benefits commensurate with their students' ability levels. *American Educational Research Journal, 28*(2), 445–480. <https://doi.org/10.3102/00028312028002445>.
- Marsh, H. W. (1992). Content specificity of relationships between academic achievement and academic self-concept. *Journal of Educational Psychology, 84*(1), 35–42. <https://doi.org/10.1037/0022-0663.84.1.35>.
- Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and revised recommendations. *Educational Psychologist, 34*(3), 155–167. https://doi.org/10.1207/s15326985ep3403_2.
- Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science, 1*(2), 133–163. <https://doi.org/10.1111/j.1745-6916.2006.00010.x>.
- *Marsh, H. W., Hau, K.-T., & Kong, C.-K. (2002). Multilevel causal ordering of academic self-concept and achievement: Influence of language of instruction (English compared with Chinese) for Hong Kong students. *American Educational Research Journal, 39*(3), 727–763. <https://doi.org/10.3102/00028312039003727>.
- Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relationships and causal ordering. *British Journal of Educational Psychology, 81*(1), 59–77. <https://doi.org/10.1348/000709910X503501>.
- *Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality and Social Psychology Bulletin, 34*(4), 542–552. <https://doi.org/10.1177/0146167207312313>.
- *Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology, 54*(2), 263–280. <https://doi.org/10.1037/dev0000393>.
- Marsh, H. W., Pekrun, R., Parker, P. D., Murayama, K., Guo, J., Dicke, T., & Arens, A. K. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology, 111*(2), 331–353. <https://doi.org/10.1037/edu0000281>.
- *Marsh, H. W., & Richards, G. E. (1986). The Rotter locus of control scale: The comparison of alternative response formats and implications for reliability, validity, and dimensionality. *Journal of Research in Personality, 20*(4), 509–528. [https://doi.org/10.1016/0092-6566\(86\)90129-7](https://doi.org/10.1016/0092-6566(86)90129-7).
- *Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*(2), 397–416. <https://doi.org/10.1111/j.1467-8624.2005.00853.x>.

- Marsh, H. W., & Yeung, A. S. (1997). Coursework selection: Relationships to academic self-concept and achievement. *American Educational Research Journal*, 34(4), 691–720. <https://doi.org/10.2307/1163354>.
- *McMillian-Robinson, M. M., Frierson, H. T., & Campbell, F. A. (2011). Do gender differences exist in the academic identification of African-American elementary-school aged children? *Journal of Black Psychology*, 37(1), 78–98. <https://doi.org/10.1177/0095798410366709>.
- *McMillian, M. M., Carr, M., Hodnett, G., & Campbell, F. A. (2016). A longitudinal study of academic identification among African American males and females. *Journal of Black Psychology*, 42(6), 508–529. <https://doi.org/10.1177/0095798415603845>.
- Meyer, J., Fleckenstein, J., & Köller, O. (2019). Expectancy value interactions and academic achievement: Differential relationships with achievement measures. *Contemporary Educational Psychology*, 58, 58–74. <https://doi.org/10.1016/j.cedpsych.2019.01.006>.
- *Mikami, A. Y., & Hinshaw, S. P. (2006). Resilient adolescent adjustment among girls: Buffers of childhood peer rejection and attention-deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology*, 34(6), 823–837. <https://doi.org/10.1007/s10802-006-9062-7>.
- Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, 120(3), 544–560. <https://doi.org/10.1037/a0032459>.
- Möller, J., Pohlmann, B., Köller, O., & Marsh, H. W. (2009). A meta-analytic path analysis of the internal/external frame of reference model of academic achievement and academic self-concept. *Review of Educational Research*, 79(3), 1129–1167. <https://doi.org/10.3102/0034654309337522>.
- Möller, J., Retelsdorf, J., Köller, O., & Marsh, H. W. (2011). The reciprocal internal/external frame of reference model: An integration of models of relations between academic achievement and self-concept. *American Educational Research Journal*, 48(6), 1315–1346. <https://doi.org/10.3102/0002831211419649>.
- *Möller, J., Zimmermann, F., & Köller, O. (2014). The reciprocal internal/external frame of reference model using grades and test scores. *British Journal of Educational Psychology*, 84(4), 591–611. <https://doi.org/10.1111/bjep.12047>.
- *Murdock, T. B., Anderman, L. H., & Hodge, S. A. (2000). Middle-grade predictors of students' motivation and behavior in high school. *Journal of Adolescent Research*, 15(3), 327–351. <https://doi.org/10.1177/0743558400153002>.
- Nagengast, B., Trautwein, U., Kelava, A., & Lüdtke, O. (2013). Synergistic effects of expectancy and value on homework engagement: The case for a within-person perspective. *Multivariate Behavioral Research*, 48, 428–460. <https://doi.org/10.1080/00273171.2013.775060>.
- *Newman, R. S. (1984). Children's achievement and self-evaluations in mathematics: A longitudinal study. *Journal of Educational Psychology*, 76(5), 857–873. <https://doi.org/10.1037/0022-0663.76.5.857>.
- Niepel, C., Brunner, M., & Preckel, F. (2014). Achievement goals, academic self-concept, and school grades in mathematics: Longitudinal reciprocal relationships in above average ability secondary school students. *Contemporary Educational Psychology*, 39(4), 301–313. <https://doi.org/10.1016/j.cedpsych.2014.07.002>.
- Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. In R. Riding & S. Rayner (Eds.), *Perception* (pp. 239–266). Ablex Publishing.
- Pietsch, J., Walker, R., & Chapman, E. (2003). The relationship among self-concept, self-efficacy, and performance in mathematics during secondary school. *Journal of Educational Psychology*, 95(95), 589–603. <https://doi.org/10.1037/0022-0663.95.3.589>.
- Pinxten, M., De Fraine, B., Van Damme, J., & D'haenens, E. (2010). Causal ordering of academic self-concept and achievement: Effects of type of achievement measure. *British Journal of Educational Psychology*, 80(4), 689–709. <https://doi.org/10.1348/000709910X493071>.
- *Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., & Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived mathematics effort expenditure. *British Journal of Educational Psychology*, 84(1), 152–174. <https://doi.org/10.1111/bjep.12028>.
- Prast, E. J., Van de Weijer-Bergsma, E., Miočević, M., Kroesbergen, E. H., & Van Luit, J. E. (2018). Relationships between mathematics achievement and motivation in students of diverse achievement levels. *Contemporary Educational Psychology*, 55, 84–96. <https://doi.org/10.1016/j.cedpsych.2018.08.002>.
- *Preckel, F., Niepel, C., Schneider, M., & Brunner, M. (2013). Self-concept in adolescence: A longitudinal study on reciprocal effects of self-perceptions in academic and social domains. *Journal of Adolescence*, 36(6), 1165–1175. <https://doi.org/10.1016/j.adolescence.2013.09.001>.
- *Quirk, M., Schwanenflugel, P. J., & Webb, M. Y. (2009). A short-term longitudinal study of the relationship between motivation to read and reading fluency skill in second grade. *Journal of Literacy Research*, 41(2), 196–227. <https://doi.org/10.1080/10862960902908467>.
- *Retelsdorf, J., Köller, O., & Möller, J. (2014). Reading achievement and reading self-concept - Testing the reciprocal effects model. *Learning and Instruction*, 29(4), 21–30. <https://doi.org/10.1016/j.learninstruc.2013.07.004>.

- *Roebers, C. M., Cimeli, P., Röthlisberger, M., & Neuenschwander, R. (2012). Executive functioning, meta-cognition, and self-perceived competence in elementary school children: An explorative study on their interrelationships and their role for school achievement. *Metacognition and Learning*, 7(3), 151–173. <https://doi.org/10.1007/s11409-012-9089-9>.
- Seaton, M., Marsh, H. W., Parker, P. D., Craven, R. G., & Yeung, A. S. (2015). The reciprocal effects model revisited: Extending its reach to gifted students attending academically selective schools. *The Gifted Child Quarterly*, 59(3), 143–156. <https://doi.org/10.1177/0016986215583870>.
- *Seaton, M., Parker, P., Marsh, H. W., Craven, R. G., & Yeung, A. S. (2014). The reciprocal relationships between self-concept, motivation and achievement: Juxtaposing academic self-concept and achievement goal orientations for mathematics success. *Educational Psychology*, 34(1), 49–72. <https://doi.org/10.1080/01443410.2013.825232>.
- Sedgwick, P. (2013). Convenience sampling. *BMJ*, 347(f6304), f6304.
- Sewasew, D., & Schroeders, U. (2019). The developmental interplay of academic self-concept and achievement within and across domains among primary school students. *Contemporary Educational Psychology*, 58, 204–212. <https://doi.org/10.1016/j.cedpsych.2019.03.009>.
- 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>.
- Sherman, D. K. (2013). Self-affirmation: Understanding the effects. *Social and Personality Psychology Compass*, 7(11), 834–845. <https://doi.org/10.1111/spc3.12072>.
- Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. *Advances in Experimental Social Psychology*, 38, 183–242. [https://doi.org/10.1016/S0065-2601\(06\)38004-5](https://doi.org/10.1016/S0065-2601(06)38004-5).
- *Sink, C. A., Barnett, J. E., & Pool, B. A. (1993). Perceptions of scholastic competence in relationship to middle-school achievement. *Perceptual and Motor Skills*, 76(2), 471–478. <https://doi.org/10.2466/pms.1993.76.2.471>.
- Skaalvik, E. M., & Hagtvet, K. A. (1990). Academic achievement and self-concept: An analysis of causal predominance in a developmental perspective. *Journal of Personality and Social Psychology*, 58(2), 292–307. <https://doi.org/10.1037/0022-3514.58.2.292>.
- Skaalvik, E. M., & Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: Relationships with teacher engagement, job satisfaction, and emotional exhaustion. *Psychological Reports*, 114(1), 68–77. <https://doi.org/10.2466/14.02.PR0.114k14w0>.
- *Skaalvik, E. M., & Valås, H. (1999). Relationships among achievement, self-concept, and motivation in mathematics and language arts: A longitudinal study. *The Journal of Experimental Education*, 67(2), 135–149. <https://doi.org/10.1080/00220979909598349>.
- Sowislo, J. F., & Orth, U. (2013). Does low self-esteem predict depression and anxiety? A meta-analysis of longitudinal studies. *Psychological Bulletin*, 139(1), 213–240. <https://doi.org/10.1037/a0028931>.
- Spinath, B., & Spinath, F. M. (2005). Development of self-perceived ability in elementary school: The role of parents' perceptions, teacher evaluations, and intelligence. *Cognitive Development*, 20(2), 190–204. <https://doi.org/10.1016/j.cogdev.2005.01.001>.
- *Stäbler, F., Dumont, H., Becker, M., & Baumert, J. (2017). What happens to the fish's achievement in a little pond? A simultaneous analysis of class-average achievement effects on achievement and academic self-concept. *Journal of Educational Psychology*, 109(2), 191–207. <https://doi.org/10.1037/edu0000135>.
- *Sticca, F., Goetz, T., Nett, U. E., Hubbard, K., & Haag, L. (2017). Short-and long-term effects of over-reporting of grades on academic self-concept and achievement. *Journal of Educational Psychology*, 109(6), 842–854. <https://doi.org/10.1037/edu0000174>.
- Stipek, D., & Iver, D. M. (1989). Developmental change in children's assessment of intellectual competence. *Child Development*, 60(3), 521–538. <https://doi.org/10.1111/1467-8624.ep7252701>.
- Stodolsky, S. S., & Grossman, P. L. (1995). The impact of subject matter on curricular activity: An analysis of five academic subjects. *American Educational Research Journal*, 32(2), 227–249. <https://doi.org/10.2307/1163430>.
- Susperreguy, M. I., Davis-Kean, P. E., Duckworth, K., & Chen, M. (2017). Self-concept predicts academic achievement across levels of the achievement distribution: Domain specificity for mathematics and reading. *Child Development*, 89(6), 2196–2214. <https://doi.org/10.1111/cdev.12924>.
- Swann Jr., 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>.
- *Szumski, G., & Karwowski, M. (2019). Exploring the Pygmalion effect: The role of teacher expectations, academic self-concept, and class context in students' math achievement. *Contemporary Educational Psychology*, 59, 1–10. <https://doi.org/10.1016/j.cedpsych.2019.101787>.

- Taris, T. W., & Kompier, M. A. J. (2014). Cause and effect: Optimizing the designs of longitudinal studies in occupational health psychology. *Work and Stress*, 28(1), 1–8. <https://doi.org/10.1080/02678373.2014.878494>.
- *Upadaya, K., & Eccles, J. (2015). Do teachers' perceptions of children's mathematics and reading related ability and effort predict children's self-concept of ability in mathematics and reading? *Educational Psychology*, 35(1), 110–127. <https://doi.org/10.1080/01443410.2014.915927>.
- Valentine, J. C., Dubois, D. L., & Cooper, H. (2004). The relationship between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, 39(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3.
- Viechtbauer, W., & Cheung, M. W. L. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2), 112–125. <https://doi.org/10.1002/jrsm.11>.
- Viljaranta, J., Tolvanen, A., Aunola, K., & Nurmi, J. E. (2014). The developmental dynamics between interest, self-concept of ability, and academic performance. *Scandinavian Journal of Educational Research*, 58(6), 734–756. <https://doi.org/10.1080/00313831.2014.904419>.
- *Walgermo, B. R., Foldnes, N., Uppstad, P. H., & Solheim, O. J. (2018). Developmental dynamics of early reading skill, literacy interest and readers' self-concept within the first year of formal schooling. *Reading and Writing*, 31(6), 1379–1399. <https://doi.org/10.1007/s11145-018-9843-8>.
- *Weidinger, A. F., Steinmayr, R., & Spinath, B. (2018). Changes in the relationship between competence beliefs and achievement in mathematics across elementary school years. *Child Development*, 89(2): e138–e156. <https://doi.org/10.1111/cdev.12806>.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*, 25(1), 68–81. <https://doi.org/10.1006/ceps.1999.1015>.
- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbretton, A. J., Freedman-Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*, 89(3), 451–469. <https://doi.org/10.1037/0022-0663.89.3.451>.
- Wigfield, A., & Karpathian, M. (1991). Who am I and what can I do? Children's self-concepts and motivation in achievement situations. *Educational Psychologist*, 26(3-4), 233–261. https://doi.org/10.1207/s15326985ep2603&4_3.
- *Wouters, S., De Fraine, B., Colpin, H., Van Damme, J., & Verschueren, K. (2012). The effect of track changes on the development of academic self-concept in high school: A dynamic test of the big-fish-little-pond effect. *Journal of Educational Psychology*, 104(3), 793–805. <https://doi.org/10.1037/a0027732>.
- Wylie, R. C. (1974). *The self-concept: theory and research on selected topics* (Vol. 2): U of Nebraska Press.
- *Yeung, A. S., & Lee, F. L. (1999). Self-concept of high school students in china: Confirmatory factor analysis of longitudinal data. *Educational and Psychological Measurement*, 59(3), 431–450. <https://doi.org/10.1177/00131649921969965>.
- *Yin, L. P. (1999). *A longitudinal study of Hong Kong Chinese university students' academic causal attributions, self-concept, learning approaches, and their causal effects on achievement*. Paper presented at the Annual Meeting of the American Educational Research Association.
- *Yoon, K. S. (1996). *Testing reciprocal causal relationships among expectancy, value and academic achievement of early adolescents: A longitudinal study*. University of Michigan.
- *Zanobini, M., & Usai, M. C. (2002). Domain-specific self-concept and achievement motivation in the transition from primary to low middle school. *Educational Psychology*, 22(2), 203–217. <https://doi.org/10.1080/01443410120115265>.
- Zelinski, E. M., Burnight, K. P., & Lane, C. J. (2001). The relationship between subjective and objective memory in the oldest old: Comparisons of findings from a representative and a convenience sample. *Journal of Aging and Health*, 13(2), 248–266. <https://doi.org/10.1177/089826430101300205>.
- *Zhou, Y.-X., Ou, C.-Q., Zhao, Z.-T., Wan, C.-S., Guo, C., Li, L., & Chen, P.-Y. (2015). The impact of self-concept and college involvement on the first-year success of medical students in China. *Advances in Health Sciences Education*, 20(1), 163–179. <https://doi.org/10.1007/s10459-014-9515-7>.
- *Zhu, M., Urhahne, D., & Rubie-Davies, C. M. (2018). The longitudinal effects of teacher judgement and different teacher treatment on students' academic outcomes. *Educational Psychology*, 38(5), 648–668. <https://doi.org/10.1080/01443410.2017.1412399>.