Journal of Personality and Social Psychology

Testing the Bottom-Up and Top-Down Models of Self-Esteem: A Meta-Analysis of Longitudinal Studies

Laura C. Dapp, Samantha Krauss, and Ulrich Orth

Online First Publication, November 3, 2022. https://dx.doi.org/10.1037/pspp0000444

CITATION
Dapp, L. C., Krauss, S., & Orth, U. (2022, November 3). Testing the Bottom-Up and Top-Down Models of Self-Esteem: A Meta-Analysis of Longitudinal Studies. *Journal of Personality and Social Psychology*. Advance online publication. https://dx.doi.org/10.1037/pspp0000444

Journal of Personality and Social Psychology:
Personality Processes and Individual Differences

© 2022 American Psychological Association
ISSN: 0022-3514

https://doi.org/10.1037/pspp0000444

Testing the Bottom-Up and Top-Down Models of Self-Esteem: A Meta-Analysis of Longitudinal Studies

Laura C. Dapp, Samantha Krauss, and Ulrich Orth
Department of Psychology, University of Bern

The present meta-analysis tests the bottom-up and top-down models of self-esteem, by synthesizing the available longitudinal evidence on prospective effects between global and domain-specific self-esteem. The bottom-up model assumes that people's domain-specific self-esteem influences their global self-esteem, whereas the top-down model assumes the reverse direction of effects. Eight domains of self-esteem were assessed: academic abilities, physical appearance, athletic abilities, morality, romantic relationships, social acceptance, mathematics, and verbal abilities. We conducted a comprehensive search of the literature, which led to the inclusion of data from 43 independent samples (total $N = 24{,}668$). One-stage meta-analytic structural equation modeling was used to estimate the coefficients of interest. There was no evidence of publication bias. Overall, the results indicated a pattern of reciprocal prospective effects between global and domain-specific self-esteem. Bottom-up effects were significant in all domains except verbal abilities (mean effect sizes ranged from .05 to .19). Top-down effects were significant in all domains except mathematics (mean effect sizes ranged from .05 to .12, except .01 in the mathematics domain). None of the moderators tested (i.e., age, gender, measure, time lag, and publication year) was significant in any of the domains, which strengthens the generalizability of the results. In sum, the findings provide support for both bottom-up and top-down effects, suggesting a reciprocal relation model between global and domain-specific self-esteem. The discussion addresses the implications of the findings for research in the field of self-esteem.

Keywords: self-esteem, self-concept, longitudinal, meta-analysis

Supplemental materials: https://doi.org/10.1037/pspp0000444.supp

In a seminal article, Shavelson et al. (1976) suggested a hierarchical model of self-esteem. According to this model, global self-esteem (i.e., individuals' overall evaluation of their worth as a person) is positioned at the apex and domain-specific self-esteem (i.e., individuals' self-evaluation in specific domains, such as academic competences, physical appearance, and social relationships) is positioned at the bottom of the model. As we will review in more detail below, research clearly indicates that global and domain-specific self-esteem are related. Two models have been proposed to explain the nature of this relation. The bottom-up model hypothesizes that domain-specific self-esteem influences people's level of global self-esteem, whereas the top-down model assumes that global self-esteem influences people's self-esteem in specific domains (e.g., Marsh & Yeung, 1998; Rentzsch & Schröder-Abé, 2022; Trautwein et al., 2006).

To date, however, only few studies have empirically tested the bottom-up and top-down models, and the inconsistent pattern of findings has not yet led to agreement about the direction of effects between global and domain-specific self-esteem. Therefore, the aim of the present meta-analysis was to comprehensively synthesize the available longitudinal data on the relation between global and domain-specific self-esteem. In the analyses, we focused on eight central domains of self-esteem: academic abilities, physical appearance, athletic abilities, morality, romantic relationships, social acceptance, and mathematics and verbal abilities.

The issue of the present research is important for the field of self-esteem for several reasons. First, both global and domain-specific self-esteem are key components of an individual's self (Vohs & Baumeister, 2012). Thus, it is essential to gain more insight into whether global and domain-specific evaluations of the self influence each other, whether there are effects in both directions of the relation, and how small or large these effects are. Although the present research is based on nonexperimental longitudinal data, researchers increasingly recognize that careful and systematic

Laura C. Dapp https://orcid.org/0000-0003-3985-6976
Samantha Krauss https://orcid.org/0000-0003-0124-347X
Ulrich Orth https://orcid.org/0000-0002-4795-515X

The results of this research have been presented at the 2021 Biennial Meeting of the Association for Research in Personality, the 2021 Virtual Convention of the Association for Psychological Science, and the 2022 Annual Convention of the Society for Personality and Social Psychology.

Data, materials, and code are available on the Open Science Framework (https://osf.io/beu29).

Laura C. Dapp played lead role in formal analysis, investigation, project administration, and writing of original draft and equal role in conceptualization. Samantha Krauss played supporting role in formal analysis, investigation, and writing of original draft. Ulrich Orth played supporting role in writing of original draft and equal role in conceptualization.

Correspondence concerning this article should be addressed to Laura C. Dapp, Department of Psychology, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland. Email: laura.dapp@unibe.ch

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

longitudinal analyses provide important information, even when the motivating research question concerns causal hypotheses (Diener et al., 2022; Grosz et al., 2020). Second, understanding the nature of the relation between global and domain-specific self-esteem has implications for research on the consequences of self-esteem. For example, research suggests that global self-esteem has benefits for many life outcomes (Orth & Robins, 2022), but if global self-esteem influences domain-specific self-esteem, it is possible that changes in domain-specific self-esteem account for (i.e., mediate) some of the positive effects of global self-esteem. Third, if domain-specific self-esteem influences global self-esteem, this knowledge could help to better understand how global self-esteem is influenced by people's experiences in important life domains, such as social relationships, school, and work (Filosa et al., 2022; Harris & Orth, 2020; Krauss & Orth, 2021; Reitz et al., 2022; Wagner et al., 2018). Thus, it is possible that domain-specific self-esteem mediates some of the effects of life experiences on global self-esteem. Fourth, evidence on the validity of the bottom-up and top-down models may contribute to research on how global and domain-specific self-esteem first emerge, when children develop their self-concept in the first years of life. For example, if the present research supports the bottom-up model but not the top-down model, this would suggest that the development of domain-specific self-esteem in early childhood might form the basis for the emergence of global self-esteem (Harter, 2003).

Theoretical Perspectives on the Relation Between Global and Domain-Specific Self-Esteem

Since the publication of Shavelson et al.'s (1976) model, there has been a long research tradition according to which self-esteem is considered a multidimensional construct that consists of hierarchically structured self-evaluations (e.g., Byrne & Shavelson, 1996; Marsh & Redmayne, 1994; Marsh & Shavelson, 1985; Shavelson & Bolus, 1982; for partial support, see Rentzsch et al., 2016), including global self-esteem and multiple domain-specific facets. Although interrelated, the various facets of domain-specific self-esteem are considered as distinct constructs (Byrne, 1996b). The hierarchical structure, in addition, implies that the facets of self-esteem are organized according to their level of specificity versus generality. More precisely, Shavelson et al. (1976) suggested that self-esteem is subdivided into specific domains, including the academic, social, and physical domain. Moreover, these domains can be further divided into subdomains. For example, the academic domain includes self-esteem related to mathematical and verbal abilities, the social domain includes self-esteem related to peer relationships and romantic relationships, and the physical domain includes self-esteem related to athletic abilities and physical appearance.

Although there is wide agreement that global and domain-specific self-esteem are interrelated, the nature of the relation remains unclear. Basically, there are two competing theoretical models, implying either that self-esteem in specific domains affects the individual's global self-esteem (bottom-up model) or that global self-esteem influences the individual's self-esteem in specific domains (top-down model).

Overall, the bottom-up model has received more attention by researchers compared to the top-down model (Brown & Marshall, 2006). First, the importance of bottom-up processes is emphasized in Shavelson et al.'s (1976) model. Although Shavelson and colleagues did not explicitly label their model as bottom-up model, they assumed that changes in self-esteem at higher levels of the hierarchy require changes in self-esteem at lower levels. Consequently, their model is often interpreted as theoretical support for the bottom-up hypothesis (Marsh & Yeung, 1998; Trautwein et al., 2006).

Also, from a developmental perspective, many researchers have emphasized the importance of bottom-up processes. For example, (Harter, 1998, 2003) suggested that global self-esteem results from an individual's accomplishments in important domains and from domain-specific feedback received by significant others. Moreover, Harter (2003) argued that children at the age of 4 or 5 years are already able to provide judgments about their competences in specific domains, such as sports, counting, drawing, or peer relationships, whereas the ability to formulate judgments about one's worth as a person (i.e., global self-esteem) does not emerge until about Age 8. The fact that the ability to report judgments of domain-specific self-esteem emerges earlier in childhood than the ability to report judgments of global self-esteem, could be interpreted as support for the bottom-up model of self-esteem.

Finally, research investigating how self-esteem is affected by objective variables, such as grades in school or physical activity, often builds on the idea of the bottom-up model. For example, Byrne and Gavin (1996; see also Byrne, 1996a) assumed that academic achievements (as reflected by grades) influence children's self-esteem with regard to specific subject matters such as mathematics and languages, which in turn influences their general academic self-esteem, which then in turn affects their level of global self-esteem. Similarly, researchers have tested whether physical activity leads to changes in self-evaluations of physical abilities and physical appearance, and whether, in turn, self-evaluations in the physical domain affect global self-esteem (Fox & Corbin, 1989; Sonstroem & Morgan, 1989).

Although many authors have focused on bottom-up processes, the top-down model of self-esteem has also received a fair amount of attention in the literature (e.g., Brown & Dutton, 1995; Brown & Marshall, 2006; Brown et al., 2001). For example, according to Brown's (1993) affective model of self-esteem development, global self-esteem develops early in life and is formed through relational and temperamental factors. This model assumes that once global self-esteem has emerged in a child, it remains relatively stable and influences the child's self-perception in specific domains. Moreover, Brown et al. (2001) proposed that high global self-esteem acts as an amplifier for positive external feedback and as a buffer against negative external feedback, and that individuals with high global self-esteem are motivated to promote and protect feelings of self-worth by attributing positive qualities to the self in specific domains. In a series of experiments, Brown et al. (2001) investigated how participants felt about themselves after experiencing manipulated failure or success. Their results showed that individuals with high global self-esteem were more prone to build and preserve positive domain-specific self-evaluations after experiencing failure than individuals with low global self-esteem (e.g., Brown et al., 2001). The results were interpreted in favor of the top-down model of self-esteem. Also, it should be noted that the hierarchical model of self-esteem is typically conceptualized as a higher order factor-analytic model. Generally, these models include directional paths that lead from the higher order to the lower order factors, consistent with the top-down model of self-esteem.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

Finally, it is important to note that many researchers—regardless of whether they focused on the bottom-up or top-down model—did not exclude the possibility of effects running in the opposite direction than stressed by their perspective (e.g., Brown et al., 2001; Harter, 1998; Shavelson & Bolus, 1982). For example, Shavelson and Bolus (1982) noted that bottom-up and top-down models could not yet be distinguished. Moreover, Marsh (1987) concluded that the hierarchical structure of self-esteem is relatively weak and therefore suggested that "it is likely that the relations are reciprocal such that the hierarchical general self-concept has some influence on specific facets and specific facets have some impact on the hierarchical construct" (p. 34). Thus, it is possible that both directions of effects operate simultaneously and that a reciprocal relation model could explain the relations between global and domain-specific self-esteem.

Empirical Evidence on the Relation Between Global and Domain-Specific Self-Esteem

Cross-sectional research clearly indicates that global and domain-specific self-esteem are related to each other (Donnellan et al., 2007; Esnaola et al., 2020; Harris et al., 2018; Marsh, 1986; von Soest et al., 2016). Although global self-esteem is positively correlated to all domains, the strength of association varies across domains. For example, in a study by Marsh (1986), correlations between global and domain-specific self-esteem ranged from .06 to .60, with most correlations falling in the range of .30–.50. Across studies, the strongest correlations are typically reported for the domain of physical appearance, whereas the weakest correlations are found for the domain of athletic abilities. The correlations with self-esteem in the domains of academic abilities and social competence are typically of medium size. Using structural equation modeling of cross-sectional data, Byrne and Gavin (1996) were the first (we are aware of) who tested the bottom-up and top-down models of self-esteem, concluding that none of the models were empirically superior to the other. Clearly, however, cross-sectional data do not allow for any conclusions about the hypothesized bottom-up and top-down effects.

Few longitudinal studies have tested for both directions of effects in multiple domains of self-esteem. For example, Rentzsch and Schröder-Abé (2022, see also Rentzsch & Schröder-Abé, 2018) examined global self-esteem and domain-specific self-esteem in five domains across four waves. Overall, their results were in favor of the top-down model, showing significant top-down effects in the domains of social criticism and performance. Marsh and Yeung (1998) investigated bottom-up and top-down effects in the academic and physical domains. Specifically, the authors tested for effects between general academic self-esteem and self-evaluations in 14 academic subdomains, as well as effects between general physical self-esteem and self-evaluations in nine physical subdomains. Only few bottom-up effects were significant. However, half of the top-down effects were significant in the academic domain and all of the top-down effects were significant in the physical domain. Thus, the findings by Marsh and Yeung (1998) overall supported the top-down model.

Several longitudinal studies examined multiple domains, but focused exclusively on bottom-up effects. For example, Donnellan et al. (2007) tested bottom-up effects in 11 domains and concluded that robust evidence emerged only in the domain of physical appearance. Similarly, von Soest et al. (2016) analyzed bottom-up effects in six domains and found evidence for significant effects only for physical appearance. Moreover, Harris et al. (2018) investigated four domains of self-esteem but did not find evidence for significant bottom-up effects.

Finally, a number of longitudinal studies tested for bottom-up and top-down effects in single domains. For example, several studies focused on the academic domain (Rosenberg et al., 1995; Trautwein, 2003; Trautwein et al., 2006), overall suggesting that bottom-up effects tended to be stronger than top-down effects. In contrast, Gogol et al. (2016) found a reverse pattern of results for the academic domain, with bottom-up effects being mostly nonsignificant and top-down effects mostly significant. Other studies focused on the physical domain. For example, Kowalski et al. (2003) reported significant bottom-up effects of athletic self-esteem on global self-esteem, but no evidence for top-down effects. Moreover, in a study by Wichstrøm and von Soest (2016), both bottom-up effects and top-down effects were significant for physical appearance self-esteem. For the social domain, Gruenenfelder-Steiger et al. (2016) found evidence for bottom-up effects but not for top-down effects, whereas in Trautwein (2003), bottom-up and top-down effects were mostly nonsignificant in this domain.

In sum, the review of studies in this field shows a relatively inconsistent picture and does not allow for reliable conclusions about bottom-up and top-down effects between global and domain-specific self-esteem. Moreover, it is possible that the effects vary across domains of self-esteem, but again the evidence does not allow for clear conclusions. In this situation, meta-analytic methods are ideally suited to gain more robust insights into the validity of the bottom-up and top-down model of global and domain-specific self-esteem.

The Present Research

The aim of the present meta-analysis was to summarize the available evidence on bottom-up and top-down effects between global and domain-specific self-esteem on the basis of longitudinal data. Table 1 gives an overview of the domains for which the effects were examined. These domains were selected for the following reasons, consistent with an earlier meta-analysis on mean-level development of domain-specific self-evaluations (Orth, Dapp, et al., 2021). The first six domains—that is, academic abilities, athletic abilities, physical appearance, social acceptance, morality, and romantic relationships—were included because they are central in theory on domain-specific self-esteem (e.g., Harter, 2015; Marsh, 1990a; Shavelson et al., 1976) as well as in key measures of the constructs, specifically, the Self-Perception Profiles developed by Harter and the Self-Description Questionnaires (SDQs) developed by Marsh (for further information, see below). Additionally, the two domains of mathematics and verbal abilities were included. Although these categories are considered subdomains of the academic self-esteem (Shavelson et al., 1976), they are explicitly distinguished in Marsh's SDQs, and they are frequently examined in the field (see e.g., Marsh & Craven, 2006, for a review). Table 1 provides information about the content of the domains, frequently used synonyms, and the corresponding subscales in the measures by Harter and Marsh. The first column of Table 1 shows brief terms used to refer to the eight domains in this meta-analysis.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

Table 1
Domains of Self-Esteem and Corresponding Subscales in Measures by Harter and Marsh

Domain	Content and synonyms	Subscales in measures by Harter	Subscales in measures by Marsh
Academic	Academic abilities, scholastic competence, intellectual abilities	PSPCSA: Cognitive Competence SPPC: Scholastic Competence SPPA: Scholastic Competence SPPCS: Scholastic Competence ASPP: Intelligence	SDQ-P: — SDQ-I: General School SDQ-II: General School SDQ-III: Academic
Appearance	Physical appearance, physical attractiveness, body satisfaction, body esteem	PSPCSA: — SPPC: Physical Appearance SPPA: Physical Appearance SPPCS: Physical Appearance ASPP: Physical Appearance	SDQ-P: Physical Appearance SDQ-I: Physical Appearance SDQ-II: Physical Appearance SDQ-III: Physical Appearance
Athletic	Athletic abilities, physical abilities, sports competences	PSPCSA: Physical Competence SPPC: Athletic Competence SPPA: Athletic Competence SPPCS: Athletic Competence ASPP: Athletic Abilities	SDQ-P: Physical Abilities SDQ-I: Physical Abilities SDQ-II: Physical Abilities SDQ-III: Physical Abilities/Sports
Morality	Morality, honesty-trustworthiness, behavioral conduct	PSPCSA: — SPPC: Behavioral Conduct SPPA: Behavioral Conduct SPPCS: Morality ASPP: Morality	SDQ-P: — SDQ-I: — SDQ-II: Honesty-Trustworthiness SDQ-III: Honesty-Trustworthiness
Romantic	Romantic relationships, romantic appeal, opposite-sex relationships, intimate relationships	PSPCSA: — SPPC: — SPPA: Romantic Appeal SPPCS: Romantic Relationships ASPP: Intimate Relationships	SDQ-P: — SDQ-I: — SDQ-II: Opposite-Sex Relations SDQ-III: Opposite-Sex Relations
Social	Social acceptance, social competence, sociability, popularity, close friendships, peer relations	PSPCSA: Peer Acceptance SPPC: Social Competence SPPA: Social Competence SPPCS: Social Acceptance ASPP: Sociability	SDQ-P: Peer Relations SDQ-I: Peer Relations SDQ-II: Same-Sex Relations SDQ-III: Same-Sex Relations
Mathematics	Mathematics abilities	PSPCSA: — SPPC: — SPPA: — SPPCS: — ASPP: —	SDQ-P: Mathematics SDQ-I: Mathematics SDQ-II: Mathematics SDQ-III: Mathematics
Verbal	Verbal abilities, language, reading, literacy	PSPCSA: — SPPC: — SPPA: — SPPCS: — ASPP: —	SDQ-P: Verbal SDQ-I: Reading SDQ-II: Verbal SDQ-III: Verbal

Note. Information on Harter's and Marsh's measures is provided in the following sources: PSPCSA = Pictorial Scale of Perceived Competence and Social Acceptance for Young Children (Harter & Pike, 1984); SPPC = Self-Perception Profile for Children (Harter, 1982, 2012b); SPPA = Self-Perception Profile for Adolescents (Harter, 2012a); SPPCS = Self-Perception Profile for College Students (Neemann & Harter, 2012); ASPP = Self-Perception Profile for Adults (Messer & Harter, 2012); SDQ-P = Self-Description Questionnaire for Preschoolers (Marsh et al., 2002); SDQ-I = Self-Description Questionnaire I (Marsh, 1992a; Marsh et al., 1984); SDQ-II = Self-Description Questionnaire II (Marsh, 1992b; Marsh et al., 2005); SDQ-III = Self-Description Questionnaire III (Marsh, 1990b; Marsh & O'Neill, 1984). Except for the PSPCSA and the SDQ-P, all measures include a scale for assessing global self-esteem. Dash indicates that no relevant subscale is included in measure.

To ensure that the results are fully comparable across the eight domains, this meta-analysis focused exclusively on the Self-Perception Profiles by Harter and the SDQs by Marsh, for the following reasons. First, these two sets of measures are the most commonly used and best validated measures for assessing multiple domains of self-esteem (Blascovich & Tomaka, 1991; Byrne, 1996b; Donnellan et al., 2015). Second, all of Harter's and Marsh's measures assess a comprehensive set of clearly defined domains (see Table 1, for an overview). Third, there is strong overlap between the domains included in Harter's and Marsh's measures, and empirical evidence suggests that the subscales that are intended to assess the same domain indeed measure similar content (i.e., show convergent validity; Donnellan et al., 2015). Consequently, focusing on these two sets of measures will allow for valid interpretation of

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

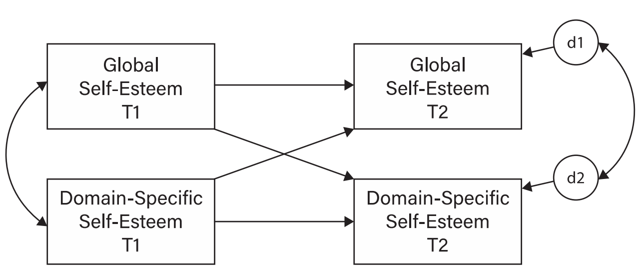
between-domain differences in the results, because methodological characteristics are held constant across domains (e.g., the general style of item wording, instructions, and response scales). Fourth, both Harter's and Marsh's measures are available in age-specific versions, providing measures of self-esteem across a broad age range from early childhood to adulthood. This characteristic was important in this meta-analysis because it allowed including samples from all developmental periods across the life span. Moreover, this characteristic also allowed to include samples from longitudinal studies that assessed the same participants in different developmental periods (e.g., with Harter's Self-Perception Profile for Children [SPPC] at Time 1 and Harter's Self-Perception Profile for Adolescents [SPPA] at Time 2). Fifth, nearly all of Harter's and Marsh's measures include a subscale measuring global self-esteem, which show strong convergent validity with the Rosenberg Self-Esteem Scale (Donnellan et al., 2015). Nevertheless, to increase the number of samples that could be included in the present meta-analysis, other measures of global self-esteem were also eligible (e.g., the Rosenberg Self-Esteem Scale; Rosenberg, 1979), as long as these measures were not a composite score of domain-specific self-esteem.[1]

In the analyses, we examined prospective effects between global and domain-specific self-esteem. To strengthen the validity of conclusions, all prospective effects were controlled for prior levels of the outcomes. Figure 1 provides a generic illustration of the effect sizes analyzed in this meta-analysis, including (a) the cross-lagged coefficients between global and domain-specific self-esteem, where the autoregressive effects of the predicted variables were controlled for (e.g., the effect of global self-esteem at Time 1 on academic self-esteem at Time 2, controlled for the effect of academic self-esteem at Time 1), (b) the stability (i.e., autoregressive) coefficients of each construct (e.g., the effect of academic self-esteem at Time 1 on academic self-esteem at Time 2), (c) the concurrent correlation between the constructs at Time 1 (e.g., the correlation between global self-esteem at Time 1 and academic self-esteem at Time 1), and (d) the residual correlation between the constructs at Time 2 (e.g., the correlation between the residual of global self-esteem at Time 2 and the residual of academic self-esteem at Time 2).

The present meta-analysis extends previous research by synthesizing the available longitudinal data on prospective effects between global and domain-specific self-esteem (no prior meta-analysis or systematic review is available on the topic). The meta-analysis adopts a comprehensive approach by examining eight domains of self-esteem (see Table 1). Importantly, all prospective effects are controlled for autoregressive effects, which significantly improves the interpretation of the findings (e.g., Cole & Maxwell, 2003; Gollob & Reichardt, 1987). Moreover, the meta-analytic method has several advantages as compared to individual studies. First, effects are estimated with high statistical power and across a heterogeneous set of samples, which increases the robustness of findings. Second, the heterogeneity of study characteristics allows for testing the effects of potential moderators (Borenstein et al., 2009; Cooper et al., 2019; Lipsey & Wilson, 2001) and thus yields insights that can hardly be provided by any individual study. Third, the aggregation of data across a heterogeneous set of studies reduces concerns about bias attributable to idiosyncrasies of individual studies. Moreover, in this meta-analysis, publication bias is unlikely, since many of the included studies did not focus specifically on the relation between global and domain-specific self-esteem, but instead, the relevant statistical information was either reported as a part of larger correlation tables in the article (or in the Supplemental Materials of the article) or we received this information upon request from the authors (for further information, see below). Taken together, the present meta-analysis advances the field by providing robust conclusions about the validity of the bottom-up and top-down models of self-esteem.

Figure 1

Generic Illustration of the Effect Sizes Analyzed



Note. Generic illustration of the effect sizes analyzed in the present meta-analysis. The model corresponds to a two-wave cross-lagged panel model, including (a) the cross-lagged effects between global self-esteem and domain-specific self-esteem, (b) the stability coefficients of each construct, (c) the concurrent correlation between the constructs at T1, and (d) the residual correlation between the constructs at T2. Residual variances (i.e., disturbances) are denoted as d1 and d2. T = time.

Method

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

Transparency and Openness

We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study, and we follow the Journal Article Reporting Standards (Appelbaum et al., 2018). Data, materials, and code are available at https://osf.io/beu29. One-stage meta-analytic structural equation modeling (MASEM) was conducted using webMASEM (Jak et al., 2021). Preliminary meta-analytic computations (i.e., Egger's regression test) were conducted with the metafor package, Version 2.4 (Viechtbauer, 2010, 2020) for R, Version 4.0.3 (R Core Team, 2020). The present research was not preregistered.

[1] In response to suggestions by reviewers, we also searched for studies using the State Self-Esteem Scale (Heatherton & Polivy, 1991; Webster et al., 2022) and the Self-Attributes Questionnaire (Pelham & Swann, 1989), using the same procedures as for the measures by Harter and Marsh. This search resulted in a total of 19 potentially relevant articles. All articles were assessed in full text. None of the studies met the inclusion criteria of the meta-analysis. Specifically, 15 studies did not include a measure of global self-esteem, one study was not empirically-quantitative, and three studies did not provide the information required for computing effect sizes. For these reasons, we continued using the restriction to measures by Harter and Marsh. However, given the evidence on convergent validity of measures of domain-specific self-esteem (Donnellan et al., 2015), we would have expected similar results if data from additional measures of domain-specific self-esteem would have been included in the present meta-analysis.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

Selection of Studies

To search for relevant studies, we used the following strategies. First, English-language journal articles, books, book chapters, and dissertations were searched in the database PsycINFO. As noted above, one of the inclusion criteria was that domain-specific self-esteem had been assessed with one of the Self-Perception Profiles by Harter or one of the SDQs by Marsh (see Table 1). Therefore, the following terms were searched in the "Test and Measures" field available in PsycINFO: *self perception profile* (to identify studies using one of Harter's measures) and *self description questionnaire* (to identify studies using one of Marsh's measures). Note that the search was unaffected by the use, or omission, of hyphens in the search terms (e.g., the term *self perception profile* also identifies entries with the term *self-perception profile*). The search was restricted to empirical-quantitative studies by using the limitation options "empirical study" and "quantitative study" available in PsycINFO. Moreover, the search was restricted to longitudinal studies based on two strategies. First, we used the limitation option "longitudinal study," which yielded 471 potentially relevant articles. Second, we omitted the limitation option "longitudinal study," but instead included *longitudinal* as a search term, which yielded 401 potentially relevant articles. After accounting for duplicates, the search resulted in a total of 592 potentially relevant studies.

To render the meta-analysis as exhaustive as possible, three additional strategies were used. First, we assessed all 103 studies that were included in a meta-analysis on mean-level change in domain-specific self-esteem (Orth, Dapp, et al., 2021). Second, we assessed 71 additional articles that were authored or coauthored by Harter or Marsh (including test manuals of the measures). Third, we assessed 21 additional articles with measures of domain-specific self-esteem of which we were aware. In sum, after accounting for overlap between all search procedures, 753 potentially relevant articles were identified.

Studies were included in the meta-analysis if the following criteria were fulfilled: (a) the study was empirically-quantitative; (b) the study used a longitudinal design (i.e., it included two or more assessments of the same sample); (c) the study included measures of global self-esteem (only continuous self-report measures were eligible), and of domain-specific self-esteem in at least one domain (only Harter's and Marsh's measures of domain-specific self-esteem were eligible); (d) the measures were unaltered across assessments (i.e., with regard to number of items, item wording, response scale), except for changes from one age-specific version of a measure to the next (e.g., from the SPPC to the SPPA); (e) enough information was given to compute the effect sizes; (f) information to compute the effect sizes was consistent and accurate across the article; (g) the sample did not undergo any intervention (e.g., if the study was an intervention study, only information from a control group that did not undergo any alternative treatment was used); (h) there was no sample overlap with other studies included in the meta-analysis (for more details, see the coding manual, which is available at https://osf.io/beu29). To decide on the eligibility of studies, all articles were assessed in full text by the first author. In addition, 60 studies were rated by the second author to obtain an estimate of interrater agreement.[2] The interrater agreement on inclusion or exclusion in the meta-analysis was high ($\kappa = 1.00$).

Of the 753 potentially relevant articles, 18 could directly be included in the meta-analysis, given that all required information was available in the article. If a study fulfilled all inclusion criteria except for providing enough information to compute effect sizes and if the article was published in 2000 or later, the corresponding author was contacted with a request for providing the required information. Based on the authors' responses, 20 additional articles could be included. In sum, the search procedures led to the inclusion of 38 articles, providing effect sizes for 43 independent samples (see Figure 2).[3]

Coding of Studies

The following data were coded: year of publication, publication type, sample size at Time 1, mean age of participants at Time 1, sample type (i.e., nationally representative, community, college students, or clinical sample), percentage of female participants, country of data collection, ethnicity, time lag between assessments, measure of global self-esteem, measure of domain-specific self-esteem, self-esteem domain, effect size information (i.e., the relevant correlation coefficients, specifically, the zero-order correlations between Time 1 global self-esteem, Time 2 global self-esteem, Time 1 domain-specific self-esteem, and Time 2 domain-specific self-esteem), and publication status of effect sizes (i.e., effect size data published in article vs. effect size data not published in article but obtained from authors).

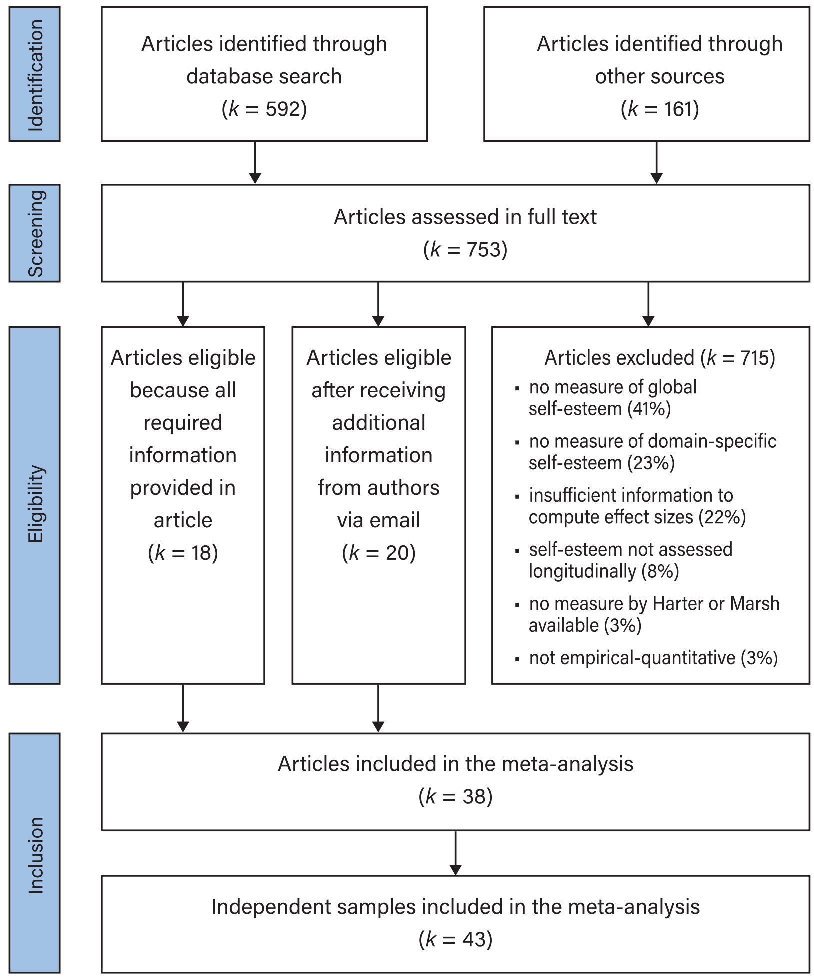
Time 1 was defined as the first assessment reported in the article at which both global and domain-specific self-esteem were measured. Time 2 was defined as the next assessment reported in the article at which at least one of the relevant variables (i.e., global and/or domain-specific self-esteem) was measured. For studies that included more than two assessments, only the first two eligible assessments were coded. If demographic information on a sample was lacking but valid information was available from other articles on the same sample, this information was used (see specific notes in Table 2). If an article did not report the mean age of participants but a valid indicator of age was available, we used this information to estimate the mean age. For example, if an age range was given (e.g., 10–12 years), the midpoint of the interval was used as estimate of mean age (e.g., 11 years).

All studies from the PsycINFO search for which the articles provided the required information ($k = 14$) were coded by two raters (i.e., the first and second author of this meta-analysis). The interrater agreement was high for categorical variables (averaged $\kappa = 1.00$) and continuous variables (averaged $r = .996$). All diverging assessments were discussed until consensus was reached. Studies for which data were obtained from the study authors (i.e., upon email request) and studies from the additional search strategies described above were coded by the first author of the present research.

[2] At the time of coding, the qualifications of the coders were as follows: The first author had a PhD in psychology and the second author had a Master's degree in psychology.

[3] Although a larger number of samples would have been desirable, most of the potentially relevant studies could not be included. The most frequent reasons for the exclusion of studies were that no measure of global self-esteem was available (41%), that no measure of domain-specific self-esteem was available (23%), and that effect size information was insufficient (22%). More detailed information on the reasons for excluding studies is provided in Figure 2. The relatively large drop from 753 potentially relevant articles to 43 included samples is also related to the fact that the search of the literature was designed to be as exhaustive as possible to ensure that the number of studies missed in the search was as small as possible. Consequently, this strategy increased the proportion of irrelevant studies included in the search results.

Figure 2
Flow Diagram of the Search and Selection Procedure



Note. The diagram has been adapted from "Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement," by D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, P. Group, and the PRISMA Group, 2009, *PLOS Medicine*, *6*(7), Article e1000097 (https://doi.org/10.1371/journal.pmed.1000097). © 2009 Moher et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. This information is also given on https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000097. See the online article for the color version of this figure.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

Meta-Analytic Procedure

First, preliminary analyses were conducted to test for publication bias. Next, weighted mean effect sizes were computed and tested for heterogeneity of effect size distributions. Last, sample, study, and measurement characteristics (i.e., age, gender, measure, time lag, and publication year) were examined as potential moderators to test for the robustness of the cross-lagged effects.

Meta-analytic computations were conducted using one-stage MASEM, which is a multivariate statistical method that allows to estimate structural equation models with meta-analytic data in a single step (Jak & Cheung, 2020; Jak et al., 2021). To test the prospective effects between global and domain-specific self-esteem, bottom-up and top-down effects were examined in eight independent models (one for each domain). Since MASEM evaluates the complete model, not only the cross-lagged effects but also the stability coefficients and the concurrent correlations between the constructs were estimated simultaneously. Consequently, the bottom-up and top-down effects were controlled for autoregressive effects of the constructs and Time 1 concurrent association between the constructs. An important advantage of MASEM is that the complete model can be estimated meta-analytically even if not all correlations are available in all primary studies. Incomplete data are handled by full information maximum likelihood (Enders, 2010; Graham, 2009).

In addition to the multivariate approach based one one-stage MASEM, we also used a univariate approach based on standardized regression coefficients. Results of this approach are reported in the

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

Table 2
Descriptive Information on the Studies Included in the Meta-Analysis

Study	Sample size	Mean age	Female (%)	Sample type	Country	Ethnicity	Measure	Time lag	ES data published
Asendorpf (2006), Study 1	312	20.30	55	College	Germany	White	Marsh	0.25	No
Asendorpf (2006)[a], Study 2	195	9.00	48	Community	Germany	White	Harter	1.00	No
Barker and Bornstein (2010), female subsample	52	10.25	100	Community	USA	White	Harter	3.59	Yes
Barker and Bornstein (2010), male subsample	78	10.25	0	Community	USA	White	Harter	3.59	Yes
Białecka-Pikul et al. (2019)	250	14.44	64	Community	Poland	White	Harter	0.80	No
Boulton et al. (2010)	115	9.50	43	Community	U.K.	White	Harter	0.42	Yes
Brendgen et al. (2007)	399	9.96	44	Community	Canada	White	Harter	1.00	No
Chen et al. (2004)	506	12.42	51	Community	China	Asian	Harter	2.00	Yes
Davidson et al. (2010)	383	12.00	44	Community	USA	White	Harter	0.50	Yes
Davison et al. (2007)	178	11.33	100	Community	USA	White	Harter	2.00	Yes
Dohnt and Tiggemann (2006)	97	6.91	100	Community	Australia	White	Harter	1.00	Yes
Donnellan et al. (2007)	409	23.26	58	Community	USA	White	Harter	2.00	No
Dufner et al. (2015)	709	11.83	54	Community	Germany	White	Harter	0.33	No
Dundas et al. (2016), control group	90	22.50	86	College	Norway	White	Marsh	0.33	No
Esnaola et al. (2020)	484	14.99	53	Community	Spain	White	Marsh	0.85	No
Fenzel (2000)	116	10.80	56	Community	USA	White	Harter	0.27	Yes
Gallagher et al. (2018)	392	15.18	52	Community	USA	White	Harter	1.00	No
Gest et al. (2005)	400	10.00	44	Community	USA	White	Harter	0.50	Yes
Harris et al. (2018)[b]	674	10.80	50	Community	USA	Hispanic	Marsh	2.00	Yes
Helgeson et al. (2007), diabetes subsample	132	12.08	53	Clinical	USA	White	Harter	1.00	No
Helgeson et al. (2007), healthy subsample	131	12.08	51	Community	USA	White	Harter	1.00	No
Kerr et al. (2009)	206	13.96	0	Community	USA	White	Harter	2.00	No
Marsh et al. (2005)	3,731	15.00	—	Community	Australia	—	Marsh	0.43	Yes
McLeod and Owens (2004)	547	10.50	50	Community	USA	Other	Harter	2.00	Yes
Mishna et al. (2014), older subsample	267	16.00	60	Community	Canada	White	Harter	1.00	No
Mishna et al. (2014), younger subsample	402	11.81	60	Community	Canada	White	Harter	1.00	No
Modecki et al. (2018)	1,146	13.29	55	Community	Australia	White	Marsh	1.00	Yes
Morin et al. (2011)	1,001	12.62	46	Community	Canada	White	Marsh	0.42	Yes
Narr et al. (2019)	169	15.21	53	Community	USA	Other	Harter	1.00	Yes
Nogueira Avelar e Silva et al. (2018)[c]	716	13.30	50	Community	Netherlands	White	Harter	0.50	No
Noordstar et al. (2016)	292	6.75	49	Community	Netherlands	White	Harter	1.00	No
O’Dea (2009)	470	12.75	63	Community	Australia	—	Harter	0.50	No
Preston et al. (2016)	130	12.00	48	Community	USA	White	Marsh	2.00	No
Putnick et al. (2020)	263	10.27	48	Community	USA	White	Harter	3.58	No
Roeder et al. (2014)	1,242	10.80	54	Community	USA	Other	Harter	0.50	Yes
Rubie-Davies and Hattie (2012)	2,298	9.64	49	Community	New Zealand	Other	Marsh	0.85	No
Schmidt et al. (2015), female subsample	198	11.90	100	Community	Switzerland	White	Harter	0.21	Yes
Schmidt et al. (2015), male subsample	230	11.90	0	Community	Switzerland	White	Harter	0.21	Yes
Siffert et al. (2012)	176	10.61	51	Community	Switzerland	White	Harter	1.00	Yes
Vanhalst et al. (2013), Study 2	526	14.95	63	Community	Belgium	White	Harter	1.00	Yes
von Soest et al. (2016)[d]	3,116	15.50	51	National	Norway	White	Harter	2.00	No
Wagnsson et al. (2014)	1,174	13.78	41	Community	Sweden	White	Harter	1.00	No
Wilson et al. (2013)	266	7.55	18	Community	USA	White	Harter	3.00	No

Note. Mean age is given in years. The column “Time lag” gives the interval between Time 1 and Time 2 in years. “ES data published” is a dichotomous variable (yes = effect size data were published in article, no = effect size data had not been published in article but were obtained from the authors). Dash indicates that information was not available. ES = effect size.
[a] Information reported in Weinert and Schneider (1999). [b] Data on morality subscale obtained from Orth et al. (2014). [c] Data on social subscale taken from Deković et al. (2018). [d] Information on the percentage of females reported in Buer et al. (2017).

Supplemental Materials (Tables S1–S4 and Figures S1–S3). Whereas one-stage MASEM has the important advantage that it allows estimating the complete model in a single step, the meta-analytic procedure based on standardized regression coefficients has other advantages, including better possibilities for conducting outlier analyses and examining funnel graphs to assess publication bias. The supplemental analyses provide useful information about whether the results replicate across the two meta-analytic approaches.

Results

Description of Studies

The meta-analytic dataset included 43 samples (Table 2 shows basic sample characteristics). Data were drawn from 38 sources, all of which were journal articles (there were no dissertations, books, or book chapters). These articles had been published between 2000 and 2020, with the median in 2012.[4]

Sample sizes ranged from 52 to 3,731 ($M = 573.7$, $SD = 764.2$, $Mdn = 312$). In sum, the samples included 24,668 participants. Participants' mean age at Time 1 ranged from 6.75 to 23.26 years ($M = 12.56$, $SD = 3.47$). The mean proportion of female participants was 53% (range = 0%–100%, $SD = 22\%$, $Mdn = 51\%$). Of the samples, 91% were community samples, 5% were samples of college students, 2% were nationally representative samples, and 2% were clinical samples. With regard to country of data collection, 42% of the samples were from the United States, 9% from Australia, 9% from Canada, 7% from Germany, 7% from Switzerland, 5% from the Netherlands, 5% from Norway, and the remaining 16% from China, Spain, New Zealand, Poland, Sweden, and United Kingdom. Taken together, almost all samples were from Western cultural contexts including the United States, Europe, Australia, and Canada, and only one sample was from Asia. No African, South American, or Central American samples were included. With regard to ethnicity, 35 samples (i.e., 81%) were predominantly White/European ("predominantly" was defined as 80% and more), one sample was predominantly Asian, one predominantly Hispanic, and six were mixed or ethnicity was unknown. Regarding the measures of domain-specific self-esteem, 79% of the studies used one of Harter's measures (44% used the SPPC, 21% used the SPPA, 9% used the SPPC at Time 1 and the SPPA at Time 2, and the remaining 5% used either the Pictorial Scale of Perceived Competence and Social Acceptance for Young Children [PSPCSA] or the Self-Perception Profile for Adults [ASPP]) and 21% used one of Marsh's measures (14% used the SDQ-II and the remaining 7% used either the SDQ-I, SDQ-III, or a combination of the SDQ measures). Regarding the measures of global self-esteem, most studies used the global self-esteem subscale included in Harter's and Marsh's measures, with the exception of four studies that used the Rosenberg Self-Esteem Scale. The time lag between Time 1 and Time 2 ranged from 0.21 to 3.59 years ($M = 1.20$, $SD = 0.92$, $Mdn = 1.00$).

Preliminary Analyses

As preliminary analyses, tests of publication bias in the cross-lagged effects were conducted. As discussed above, no pattern of publication bias in the effect sizes was expected, because many studies included in this meta-analysis focused on other research questions. Moreover, in more than 50% of the studies, the relevant statistics were not reported in the article but obtained from the studies' authors (see Method section). Two methods were used to test for publication bias. First, effect sizes that were published in the articles were compared to effect sizes that were not published (but obtained from the studies' authors upon request) by using publication status as a moderator variable in the analysis. If the size and significance of an effect size influences whether it is published or not, then this comparison should yield a significant difference between published and unpublished effect sizes. Given that 16 tests were conducted (for eight domains and both directions of the effects), the significance level was adjusted to $p < .003$, following the Bonferroni correction (i.e., dividing .05 by 16). All tests were nonsignificant (see Table 3). For the domains of romantic and mathematics self-esteem, the model did not converge, which might be due to the small number of samples (see Jak & Cheung, 2020). Second, Egger's regression test (Egger et al., 1997) was conducted based on Fisher's z-values of the correlations with the regtest function of the metafor package (Viechtbauer, 2020), including the Knapp and Hartung adjustment (Knapp & Hartung, 2003). Given that 48 tests were conducted (for eight domains and all six correlations), the significance level was adjusted to $p < .001$, following the Bonferroni correction (i.e., dividing .05 by 48). Egger's regression test was nonsignificant for all effect sizes, suggesting that there was no evidence of publication bias in the data (see Table S5 in the Supplemental Materials). Together, both methods suggested that there was no evidence of publication bias in the meta-analytic dataset.

Effect Size Analyses

For each self-esteem domain, one coherent model was tested, whereby weighted mean effect sizes were estimated for (a) the cross-lagged effects between global and domain-specific self-esteem, (b) the stability effects of global and domain-specific self-esteem, (c) the concurrent correlation between global and domain-specific self-esteem at Time 1, and (d) the residual correlation at Time 2. Table 4 reports the results of the effect size analyses and Figure 3 provides a graphical summary of the weighted mean effect sizes.

For the academic, appearance, athletic, morality, romantic, and social domain, significant cross-lagged effects emerged in the direction of both bottom-up and top-down processes. More precisely, global self-esteem positively predicted later domain-specific self-esteem and, vice versa, domain-specific self-esteem positively predicted later global self-esteem. The cross-lagged effects ranged from .05 (romantic self-esteem) to .19 (appearance self-esteem) for bottom-up effects and from .05 (athletic self-esteem) to .12 (appearance self-esteem) for top-down effects.

For the academic subdomains of mathematics and verbal abilities, the pattern of results was less consistent. For the mathematics domain, only the bottom-up effect was significant, whereas for the verbal abilities domain, only the top-down effect was significant. However, given that the number of samples that provided information on the domains of mathematics and verbal abilities was small (i.e., four samples each), the conclusions that can be drawn from these analyses are limited. Thus, although the bottom-up and top-down effects were not consistently significant in the mathematics and verbal abilities

[4] Note that articles from any year of publication were eligible for inclusion in the meta-analysis. However, none of the studies published prior to 2000 were included in the meta-analytic dataset, because none of these studies met the inclusion criteria (e.g., because the studies were not longitudinal or because information on effect sizes was insufficient).

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

Table 3

Tests of Publication Bias in Cross-Lagged Effects, Based on Comparing Studies for Which Effect Size Data Were Published in Article Versus Not Published in Article

Variable	*B*	*SE*	*p*
Academic			
D → G	.08	.06	.175
G → D	.10	.06	.110
Appearance			
D → G	−.09	.08	.251
G → D	−.04	.09	.696
Athletic			
D → G	−.03	.05	.594
G → D	.05	.06	.350
Morality			
D → G	.03	.07	.642
G → D	−.05	.06	.412
Romantic			
D → G	—	—	—
G → D	—	—	—
Social			
D → G	.04	.04	.289
G → D	.02	.04	.592
Mathematics			
D → G	—	—	—
G → D	—	—	—
Verbal			
D → G	.09	.05	.046
G → D	.07	.03	.041

Note. The differences between effect sizes from studies for which effect size data were published in article (effect size data published = 1) versus not published in article (effect size data not published = 0) were tested with moderator analyses in webMASEM. For the domains of romantic and mathematics self-esteem, the model did not converge (indicated by dashes), which might be due to the small number of samples. The significance level was adjusted to $p < .003$ (Bonferroni correction). D = domain-specific self-esteem; G = global self-esteem; *SE* = standard error; MASEM = meta-analytic structural equation modeling.

domain, the overall findings corresponded to a pattern of reciprocal effects between global and domain-specific self-esteem, suggesting that bottom-up and top-down processes operate simultaneously. Moreover, even if the bottom-up and top-down effects for mathematics and verbal abilities were not consistently significant, it is important to note that the estimates of all effects were in the expected direction (i.e., all had a positive sign).

Moderator Analyses

The findings reported in Table 4 suggested that there was little heterogeneity in effect sizes for most domains. Nevertheless, moderator analyses were conducted to test whether age (i.e., participants' mean age at Time 1), gender (i.e., proportion of female participants in the sample), measure used to assess self-esteem (i.e., Harter vs. Marsh), time lag between the assessments, and publication year moderated the size of bottom-up and top-down effects. Given that the number of samples was small for the domains of mathematics and verbal abilities (i.e., four samples each), moderator analyses were conducted only for the academic, appearance, athletic, morality, romantic, and social domain (i.e., the first six domains shown in Table 1), for reasons of statistical power (Borenstein et al., 2009; Cooper et al., 2019; Viechtbauer et al., 2015). As recommended by Jak and Cheung (2020), the five moderators were tested in separate models. The variables age (in years), gender (i.e., proportion of female participants in the sample), time lag (in years), and publication year were continuous and, following the recommendations by Jak and Cheung (2020), centered for the moderator analyses, whereas the variable measure was dichotomous and included as such. The intercorrelations between the moderators are reported in Table 5.

Table 6 shows the results of the moderator analyses. Because of the large number of tests (i.e., five moderators for each of the 12 effect sizes), the significance level was adjusted to $p < .001$, following the Bonferroni correction (i.e., dividing .05 by 60). The results indicated that none of the moderators were significant, except for the moderator effect of measure on the bottom-up effect in the romantic domain. Thus, the findings suggest that the bottom-up and top-down effects held across samples varying with regard to age, gender, measure, time lag, and publication year, which strengthens the generalizability of the results.

Discussion

In this meta-analysis, we synthesized the available longitudinal evidence on the bottom-up and top-down models of self-esteem. Analyses were based on data from 43 independent samples, including 24,668 participants. Overall, the results indicated a pattern of reciprocal prospective effects between global and domain-specific self-esteem. For the academic, appearance, athletic, morality, romantic, and social domain, both the bottom-up and top-down effects were significant, whereas in the domains of mathematics and verbal abilities, the pattern of findings was less consistent (for the mathematics domain, only the bottom-up effect was significant, whereas for the verbal abilities domain, only the top-down effect was significant). In sum, the findings supported both bottom-up and top-down effects (with the exception of mathematics and verbal abilities), suggesting a reciprocal relation model between global and domain-specific self-esteem. To assess the robustness of the effects, we tested whether age, proportion of gender, type of measure, time lag between assessments, and publication year moderated the findings. However, none of the moderators was significant (except for one moderator effect of measure in the romantic domain), which strengthens the generalizability of the results.

Implications of the Findings

As reviewed in the introduction section, the hierarchical model of self-esteem by Shavelson et al. (1976) led to the proposal of two competing models of the relation between global and domain-specific self-esteem, that is, the bottom-up model (e.g., Harter, 2003) and the top-down model (e.g., Brown et al., 2001). However, the empirical evidence from research in this field had been inconsistent and did not yet allow for clear conclusions about the validity of the models (for key studies in this field, see Donnellan et al., 2007; Marsh & Yeung, 1998; Rentzsch & Schröder-Abé, 2022; Trautwein et al., 2006; von Soest et al., 2016). Based on the present meta-analysis, which synthesized the longitudinal evidence from 43 samples, it is now possible to draw relatively robust conclusions. For most domains, the findings suggested that both bottom-up and top-down effects are significant and of about similar size.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

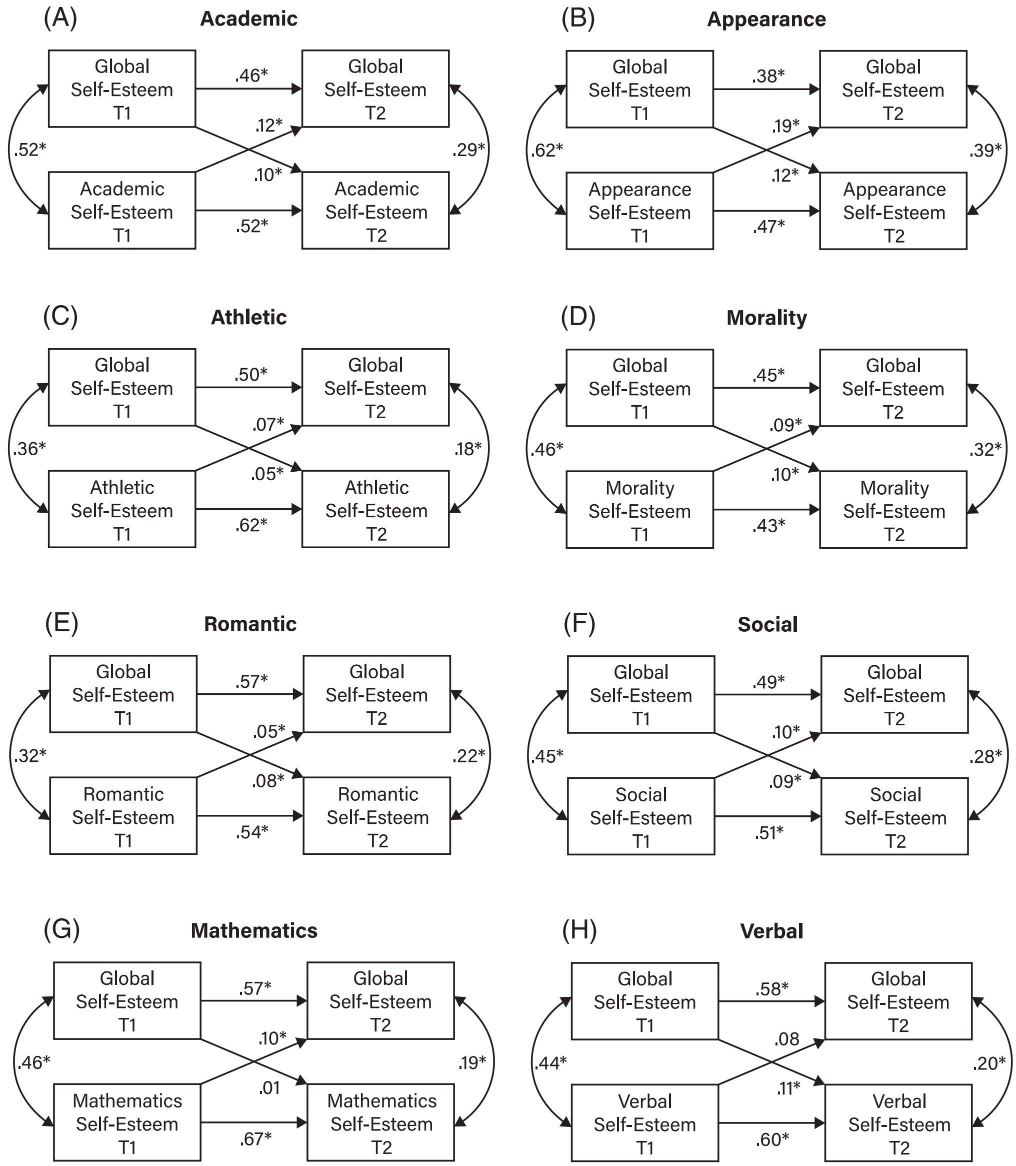
Table 4
Summary of Effect Sizes for Relations Between Global and Domain-Specific Self-Esteem

Variable	k	N	Estimate	SE	p	Heterogeneity[a]
Academic						
$r_{G,D\ T1}$	26	17,534	.52*	.03	<.001	.02
D → G	26	17,534	.12*	.03	<.001	.01
G → D	26	17,534	.10*	.03	.004	.01
G → G	26	17,534	.46*	.03	<.001	.01
D → D	26	17,534	.52*	.03	<.001	.01
$r_{G,D\ T2}$	26	17,534	.29*	.03	<.001	.01
Appearance						
$r_{G,D\ T1}$	25	15,118	.62*	.02	<.001	.01
D → G	25	15,118	.19*	.04	<.001	.01
G → D	23	14,855	.12*	.05	.008	.01
G → G	25	15,118	.38*	.04	<.001	.01
D → D	23	14,855	.47*	.05	<.001	.02
$r_{G,D\ T2}$	23	14,855	.39*	.02	<.001	.01
Athletic						
$r_{G,D\ T1}$	17	12,084	.36*	.02	<.001	.00
D → G	17	12,084	.07*	.02	<.001	.00
G → D	17	12,084	.05*	.02	.020	.00
G → G	17	12,084	.50*	.03	<.001	.01
D → D	17	12,084	.62*	.04	<.001	.02
$r_{G,D\ T2}$	17	12,084	.18*	.01	<.001	.00
Morality						
$r_{G,D\ T1}$	15	8,050	.46*	.03	<.001	.01
D → G	15	8,050	.09*	.03	.008	.00
G → D	15	8,050	.10*	.03	<.001	.00
G → G	15	8,050	.45*	.04	<.001	.01
D → D	15	8,050	.43*	.03	<.001	.01
$r_{G,D\ T2}$	15	8,050	.32*	.02	<.001	.01
Romantic						
$r_{G,D\ T1}$	11	10,235	.32*	.02	<.001	.00
D → G	11	10,235	.05*	.02	.014	.00
G → D	11	10,235	.08*	.02	<.001	.00
G → G	11	10,235	.57*	.02	<.001	.00
D → D	11	10,235	.54*	.04	<.001	.01
$r_{G,D\ T2}$	11	10,235	.22*	.02	<.001	.00
Social						
$r_{G,D\ T1}$	32	20,174	.45*	.02	<.001	.01
D → G	32	20,174	.10*	.02	<.001	.00
G → D	31	19,791	.09*	.02	<.001	.00
G → G	32	20,174	.49*	.03	<.001	.01
D → D	31	19,791	.51*	.03	<.001	.01
$r_{G,D\ T2}$	31	19,791	.28*	.02	<.001	.01
Mathematics						
$r_{G,D\ T1}$	4	6,643	.46*	.03	<.001	.00
D → G	4	6,643	.10*	.04	.015	.00
G → D	4	6,643	.01	.05	.790	.00
G → G	4	6,643	.57*	.05	<.001	.00
D → D	4	6,643	.67*	.07	<.001	.01
$r_{G,D\ T2}$	4	6,643	.19*	.04	<.001	.00
Verbal						
$r_{G,D\ T1}$	4	6,643	.44*	.05	<.001	.01
D → G	4	6,643	.08	.06	.152	.00
G → D	4	6,643	.11*	.05	.036	.00
G → G	4	6,643	.58*	.05	<.001	.00
D → D	4	6,643	.60*	.04	<.001	.00
$r_{G,D\ T2}$	4	6,643	.20*	.05	<.001	.00

Note. The table shows meta-analytic effect size estimates for bottom-up effects (D → G), top-down effects (G → D), stability effects (G → G and D → D), concurrent correlations at Time 1 ($r_{G,D\ T1}$), and residual correlations at Time 2 ($r_{G,D\ T2}$). Computations were made with one-stage meta-analytic structural equation modeling (MASEM) in webMASEM. k = number of samples; N = total number of participants in the k samples; SE = standard error; D = domain-specific self-esteem; G = global self-esteem.
[a] When using MASEM, between-study variances do not quantify the heterogeneity of the effect size variable, but the heterogeneity of the underlying correlation matrix.
* $p < .05$.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

Figure 3
Graphical Summary of Weighted Mean Effect Sizes



Note. T = time.
* $p < .05$.

Specifically, across domains, the average bottom-up effect was .10 and the average top-down effect was .08 (based on the data reported in Table 4). When omitting the domains of mathematics and verbal abilities (for which the pattern of significance differed from the other domains), the average effects were not much altered (i.e., .10 for bottom-up effects and .09 for top-down effects). Thus, the present findings provide support for both bottom-up and top-down effects, indicating that the relation between global and domain-specific self-esteem is best described by the reciprocal relation model, as suggested by Marsh (1987).

The largest effect size emerged for the bottom-up effect of appearance self-esteem ($\beta = .19$). This finding is in line with several studies in this field, which had found the strongest, and sometimes the only, bottom-up effect for this domain (Donnellan et al., 2007; von Soest et al., 2016; Wichstrøm & von Soest, 2016). The literature suggests that two different mechanisms could account for the relatively large bottom-up effect of appearance self-esteem. One critical factor might be the emphasis that is placed on people's (in particular, girls' and women's) physical appearance in most contemporary societies (Harter, 2003). Indeed, there has been considerable support for the thin-ideal internalization theory, which suggests that due to the repeated exposure to the society's body and beauty ideals, individuals place high value on physical attributes such as thinness and attractiveness when evaluating their overall

Table 5
Intercorrelations Between Study Moderators

Variable	1	2	3	4	5
1. Mean age at Time 1	—				
2. Female (proportion)	.12	—			
3. Measure[a]	−.30*	−.05	—		
4. Time lag	−.19	−.14	.17	—	
5. Publication year	.13	−.03	−.17	.05	—

[a] 0 = Marsh, 1 = Harter.
* $p < .05$.

worth as a person (Thompson & Stice, 2001; Wichstrøm & von Soest, 2016). A second critical factor is that physical appearance is always on display for oneself and others (i.e., the observability of physical appearance is high). Consequently, individuals have much less control over whether, when, and how others perceive their adequacy in the domain of appearance, compared to other characteristics, such as academic or athletic competence.

In contrast to the domain of physical appearance, the effect sizes were smaller in the domain of athletic abilities. Thus, although both domains are part of the physical self in Shavelson et al.'s (1976) model, the meta-analytic findings suggest that self-esteem in the domain of athletic abilities is less strongly linked to changes over time in global self-esteem than is self-esteem in the domain of physical appearance. These longitudinal findings are consistent with the fact that athletic self-esteem typically shows smaller cross-sectional correlations with global self-esteem than does appearance self-esteem (e.g., Donnellan et al., 2007; Harter, 2003). In future research, it would be interesting to test why appearance self-esteem is more closely linked to global self-esteem compared to athletic self-esteem. In our opinion, the mechanisms discussed above

Table 6
Moderator Analyses for Sample, Study, and Methodological Characteristics Predicting Bottom-Up and Top-Down Effects

	Bottom-up (D → G)				Top-down (G → D)			
Moderator	*k*	*B*	*SE*	*p*	*k*	*B*	*SE*	*p*
Academic								
Mean age at Time 1	26	−.01	.01	.394	26	−.00	.01	.966
Female (proportion)	25	−.00	.00	.031	25	−.00	.00	.407
Measure[a]	26	−.11	.06	.047	26	−.14	.06	.020
Time lag	26	.03	.03	.360	26	−.03	.03	.307
Publication year	26	.00	.01	.779	25	−.00	.01	.901
Appearance								
Mean age at Time 1	25	−.00	.01	.733	23	−.01	.01	.217
Female (proportion)	24	−.00	.00	.588	22	.00	.00	.644
Measure[a]	25	.26	.08	.002	23	.28	.10	.004
Time lag	25	.02	.04	.506	23	−.03	.04	.419
Publication year	25	.00	.01	.886	23	−.01	.01	.564
Athletic								
Mean age at Time 1	17	−.01	.01	.301	17	−.00	.00	.308
Female (proportion)	16	−.00	.00	.137	16	−.00	.00	.482
Measure[a]	17	.06	.05	.227	17	−.02	.04	.640
Time lag	17	.03	.03	.188	17	−.02	.03	.530
Publication year	17	.01	.00	.065	17	−.00	.00	.984
Morality								
Mean age at Time 1	15	−.01	.01	.140	15	.00	.01	.669
Female (proportion)	14	−.00	.00	.258	14	−.00	.00	.731
Measure[a]	15	.04	.07	.542	15	.07	.06	.264
Time lag	15	.01	.03	.650	15	−.01	.02	.591
Publication year	15	−.00	.01	.642	15	−.01	.00	.191
Romantic								
Mean age at Time 1	11	.01	.01	.009	11	.00	.01	.866
Female (proportion)	10	.00	.00	.174	10	.01	.00	.168
Measure[a]	11	.12	.02	<.001	11	.08	.03	.002
Time lag	11	.04	.02	.059	11	.04	.02	.056
Publication year	11	.00	.00	.343	11	.01	.00	.039
Social								
Mean age at Time 1	32	−.01	.01	.022	31	−.00	.01	.344
Female (proportion)	31	−.00	.00	.294	30	−.00	.00	.887
Measure[a]	32	.03	.05	.532	31	.01	.04	.904
Time lag	32	.04	.02	.067	31	.01	.02	.614
Publication year	32	−.00	.00	.538	31	−.01	.00	.035

Note. Regression coefficients of moderators are unstandardized. Reported are the moderator effects for the cross-lagged effects, based on models in which the moderator effects were tested for all longitudinal effects simultaneously. For the domains of mathematics and verbal abilities, the number of studies did not provide sufficient power for testing moderators. The significance level was adjusted to $p < .001$ (Bonferroni correction). k = number of samples; D = domain-specific self-esteem; G = global self-esteem; *SE* = standard error.
[a] 0 = Marsh, 1 = Harter.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

(i.e., internalization of societal beauty ideals and differences in observability), provide promising hypotheses for explaining the differences between appearance and athletic aspects of the physical self.

For the domains of mathematics and verbal abilities, the pattern of results was less consistent. For the mathematics domain, only the bottom-up effect was significant, whereas for the verbal abilities domain, only the top-down effect was significant. As noted above, the number of samples was small for mathematics and verbal abilities, which limits the conclusions that can be drawn for these domains. Nevertheless, the following two reasons could explain why the results did not consistently support reciprocal effects for these domains. First, according to the hierarchical model by Shavelson et al. (1976), mathematics and verbal abilities are subdomains of the general academic domain. Hence, in this model, self-evaluations of specific academic subdomains such as mathematics and verbal abilities are more distant to global self-esteem compared to general academic self-esteem. Moreover, all other categories examined in the present meta-analysis (i.e., academic, appearance, athletic, morality, romantic, and social self-esteem) refer to relatively broad domains. Thus, self-esteem in the domains of mathematics and verbal abilities might be more closely related to actual behavior and performance, and therefore, might be more strongly linked to specific experiences than to global self-esteem. Second, meta-analytic evidence indicates that self-esteem in the domains of mathematics and verbal abilities shows a normative developmental trajectory that differs fundamentally from the trajectory of global self-esteem (Orth, Dapp, et al., 2021; Orth et al., 2018). More precisely, whereas mean levels of self-esteem in the domains of mathematics and verbal abilities tend to decrease during adolescence and young adulthood (Orth, Dapp, et al., 2021), mean levels of global self-esteem are constant in early adolescence and increase in later adolescence and young adulthood (Orth et al., 2018). Moreover, in contrast to mathematics and verbal abilities, general academic self-esteem shows a positive developmental trajectory similar to global self-esteem (Orth, Dapp, et al., 2021). These findings are consistent with the notion that constructs that follow divergent normative trajectories are less closely interrelated (in terms of unidirectional or reciprocal effects) compared to constructs that follow similar normative trajectories.

As noted above, the meta-analytic effects ranged from .05 to .19 (except for the top-down effect in the mathematics domain, which was .01), with average values of .10 (bottom-up effects) and .08 (top-down effects). At first sight, these values might be considered small, especially when assessed with effect size conventions for interpreting correlation coefficients (Cohen, 1992). However, conventions for correlations should not be applied to cross-lagged effects for several reasons (for a detailed discussion, see Orth et al., 2022). First, cross-lagged effects are based on longitudinal data, whereas effect size conventions for correlations typically refer to associations based on cross-sectional data. Given that most psychological constructs change over time, longitudinal associations are typically smaller than concurrent associations. Second, cross-lagged effects are statistically controlled for the stability of the constructs (Adachi & Willoughby, 2015). The stability of a construct often explains a large part of its variance over time (which is also true for global and domain-specific self-esteem, as indicated by the present meta-analytic findings), which severely restricts the theoretical range of cross-lagged effects from other constructs. Third, cross-lagged effects are also controlled for the concurrent correlation of the constructs at Time 1. Specifically, the bivariate correlation between the Time 1 predictor and the Time 2 outcome is partitioned into two components, including (a) the path consisting of the concurrent Time 1 correlation between the constructs and the Time 1–Time 2 stability path of the outcome, and (b) the direct path from the Time 1 predictor on the Time 2 outcome (Kenny, 1979). Since the bivariate correlation is equal to the sum of the two paths, it is rare that one path (i.e., the cross-lagged effect) has the same size as the bivariate correlation (in fact, the other path typically explains a large part of the bivariate correlation). For these reasons, a recent meta-analytic project established benchmarks for cross-lagged effects in several fields of psychology, including social-personality and developmental psychology (Orth et al., 2022). Based on the empirical distribution of effect sizes, the findings suggested that a cross-lagged effect of .03 should be considered as small (corresponding to the 25th percentile of the distribution), .07 as medium (50th percentile), and .12 as large (75th percentile). Thus, most of the cross-lagged effects found in this meta-analysis should be considered as of medium size. These conclusions are also supported by the fact that the present effect sizes correspond to meta-analytic estimates of well-established cross-lagged effects in other fields. For example, prospective effects between low positive emotionality and depression/anxiety range from .06 to .09 (Khazanov & Ruscio, 2016), effects between loneliness and social anxiety range from .09 to .12 (Maes et al., 2019), and effects between social support and posttraumatic stress disorder range from .09 to .10 (Wang et al., 2021).

The findings of the present research may also have implications for understanding the development of self-esteem early in life. As noted in the introduction section, researchers had built on both the bottom-up model and top-down model to derive hypotheses about how self-esteem emerges and develops in young children (see also Orth & Robins, 2019). For example, building on the bottom-up model, Harter (1998, 2003) suggested that domain-specific self-esteem develops first, based on the child's accomplishments in important domains and on feedback received by parents, teachers, and peers. Harter (2003) assumed that domain-specific self-esteem later forms the basis for the emergence of global self-esteem. In contrast, building on the top-down model, Brown (1993) suggested that global self-esteem develops first and is formed through early relational and temperamental factors. Brown (1993) assumed that global self-esteem then influences the child's self-perception in any specific domain. The present findings suggest that both theoretical perspectives on the early development of self-esteem might be valid. In fact, it is possible that the reciprocal relation model captures not only processes in middle childhood, adolescence, and adulthood, but also the emergence of self-esteem in early childhood. However, it is important to note that the meta-analytic dataset did not include samples younger than 6 years. Consequently, any conclusions about bottom-up and top-down processes in early childhood must be considered tentative, requiring further research.

The present findings have also broader implications for research in the field of self-esteem. The reciprocal relation model suggests that global and domain-specific self-esteem mutually influence each other. Consequently, it might be useful to examine both global and domain-specific self-esteem when investigating key questions of the field. However, research often focuses exclusively, or at least predominantly, on global self-esteem. For example, research suggests that low self-esteem is a risk factor for mental health problems

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

(Sowislo & Orth, 2013; Zeigler-Hill, 2011). However, the majority of these studies used measures of global rather than domain-specific self-esteem (for exceptions, see Orth et al., 2014; Steiger et al., 2014). More generally, reviews of the literature suggest that high self-esteem is beneficial in many important life domains, including social relationships, school, and work (Donnellan et al., 2011; Orth & Robins, 2022). Again, however, the majority of studies used measures of global self-esteem, although the specificity-matching principle suggests that more specific outcomes are better predicted by domain-specific self-esteem than by global self-esteem (Swann et al., 2007). Consequently, a general recommendation for the field of self-esteem is to more often examine both global and domain-specific self-esteem. Thus, it would be useful to include global and domain-specific measures of self-esteem in ongoing and planned longitudinal studies.

The present findings may also have implications for research and practice in the field of self-esteem interventions (Niveau et al., 2021; O'Mara et al., 2006). Given the empirical support for the reciprocal relation model, it is possible that interventions targeting global self-esteem also lead to improvements in domain-specific self-esteem and, vice versa, interventions targeting domain-specific self-esteem also raise global self-esteem. Thus, future research could compare the efficacy of interventions focusing on global self-esteem, domain-specific self-esteem, or both, and then select those components of self-esteem that can be changed with higher effectiveness and in a sustained way. It is important to note that self-esteem interventions should not be used broadly until there is robust knowledge about the effectiveness of the specific interventions (Brummelman & Sedikides, 2020; Orth & Robins, 2022).

Limitations, Strengths, and Future Directions

Several limitations should be considered when interpreting the findings. Although this meta-analysis was based exclusively on longitudinal studies, a limitation is that the analyses do not allow for strong causal conclusions about the relations between global and domain-specific self-esteem. As in all nonexperimental studies, it is possible that the effects are confounded by unmeasured third variables (e.g., Little et al., 2007). Nevertheless, analyzing longitudinal data is insightful because it can provide information about whether the observed effects are consistent with a causal model of the relation between the constructs.

Another limitation is that the samples predominantly came from Western cultural contexts and were of White/European ethnicity. Thus, this meta-analysis did not allow to test whether the findings hold in samples from non-Western countries or in non-White samples. In future research, it will be important to evaluate the degree to which bottom-up and top-down processes generalize across cultures and ethnicities.

In the analyses, it was not possible to test whether the effect sizes differed significantly between domains. One reason is that the one-stage MASEM analyses were statistically already relatively complex, even when the cross-lagged and stability effects were estimated for only one domain per model. Including multiple, or even all eight, domains in the same MASEM analyses would have likely overburdened the computations. Moreover, many of the included studies provided data for only few domains. In particular, for some domains (i.e., mathematics and verbal abilities), only few effect sizes were available, despite the comprehensive search of studies. Consequently, there would have been a relatively large degree of missing data across studies and domains, which would have further increased the complexity of the analyses. In future research with individual-level data, it would be worthwhile to test whether the bottom-up and top-down effects differ statistically across domains. Moreover, future research would benefit from conducting more longitudinal studies in which a comprehensive set of self-esteem domains is assessed. Nevertheless, regarding the present meta-analysis, we believe that the meta-analytic estimates of mean effect sizes and standard errors provide a good basis for evaluating and comparing the size of effects across domains.

The present meta-analytic computations were based on zero-order correlations that were not corrected for attenuation due to measurement error. Thus, it is possible that the observed between-study heterogeneity of effect sizes is partly due to the lack of correcting for unreliability. At the same time, presence of measurement error does not necessarily cause systematic bias in the size of cross-lagged effects. In a recent meta-analysis across different fields of psychology, the magnitude of cross-lagged effects did not differ significantly between studies that examined latent versus manifest construct factors (Orth et al., 2022). Nevertheless, in individual studies, it is clearly recommended to measure constructs as latent variables and to control for unreliability of the measures (Cole & Preacher, 2014). Moreover, in future research on the bottom-up and top-down models, it would be interesting to test whether the results generalize across studies based on latent versus manifest variables.

Finally, it is possible that people's beliefs about the importance versus unimportance of specific domains moderate the effects of domain-specific self-esteem on global self-esteem. Following a Jamesian perspective, some researchers have argued that the effect of domain-specific self-esteem on global self-esteem depends on the importance an individual places on each specific domain (e.g., Harter, 2003; Rosenberg et al., 1995). More precisely, the influence of a specific domain on global self-esteem might be stronger when the domain is perceived as important, and weaker when the domain is perceived as unimportant. Several studies have tested these questions using different approaches such as weighting reports of domain-specific self-esteem by the individuals' subjective importance ratings or by assigning domain-specific weights on the group or subgroup level (e.g., Donnellan et al., 2007; Hardy & Moriarty, 2006; Marsh, 1993, 1995; Pelham, 1995a, 1995b; Scalas et al., 2013). Although the hypothesis that individual differences in the subjective importance of domains influence bottom-up processes is theoretically appealing, empirical tests have provided only limited support. For example, a series of studies by Marsh (Marsh, 1986, 1993, 2008) suggested that weighting domain-specific self-esteem by its importance—either on the individual or on the group level—does not explain much incremental variance in global self-esteem, over and above the average effects of domain-specific self-esteem. Nevertheless, the available studies used relatively different methodological approaches and, consequently, it is not yet clear how robust the results on the subjective importance of specific domains are. However, given that importance ratings were not available in almost all individual studies included in this meta-analysis, we could not address this issue empirically. In future research it would be important and interesting to test these hypotheses by meta-analytic aggregation of data across a larger number of samples.

The present research also has important strengths. A crucial advantage of meta-analyses consists in the aggregation of data

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

across a large number of diverse studies with many participants, which significantly increases the robustness and generalizability of the findings. In the present research, 43 independent samples provided data from more than 24,000 individuals. Moreover, the present research examined eight important domains of self-esteem, which provides comprehensive information on bottom-up and top-down effects across domains.

Another major strength is the longitudinal nature of the data. Specifically, testing prospective effects and controlling for autoregressive effects in the constructs significantly strengthened the validity of the conclusions. By using a multivariate meta-analytic approach, it was possible to estimate all coefficients of interest—that is, the cross-lagged effects, the stability effects, and the concurrent correlations between the constructs—simultaneously. Moreover, there was no evidence of publication bias in the meta-analytic dataset, as indicated by approaches to examine publication bias (including the comparison of published vs. unpublished effect size data).

Also, as noted above and as reported in the Supplemental Materials, we replicated the present analyses using a different meta-analytical approach. Specifically, in addition to the multivariate approach based on one-stage MASEM (Jak & Cheung, 2020; Jak et al., 2021), we used a univariate approach based on standardized regression coefficients. Whereas one-stage MASEM has the important advantage that it allows estimating the complete model in a single step, the meta-analytic procedure based on standardized regression coefficients has other advantages, including better possibilities for conducting outlier analyses and examining funnel graphs to assess publication bias. Importantly, the results of the alternative meta-analytic approach were quite similar to the results of the multivariate approach and the conclusions were essentially the same, which strengthens confidence in the findings.

We believe that there are a number of important directions for future research, some of which were already mentioned earlier. These include (a) identifying the mechanisms that explain the particularly strong relation between appearance self-esteem and global self-esteem, (b) testing whether individual differences in the subjective importance of domains moderate bottom-up and top-down processes, (c) testing whether the findings hold in samples of very young children, when subjective feelings of global and domain-specific self-esteem first emerge, (d) testing whether there are important cultural and ethnic differences in the validity of the bottom-up and top-down models of self-esteem, and (e) testing whether the findings generalize across studies using latent versus manifest variables.

An additional direction for future research is related to statistical modeling of effects between global and domain-specific self-esteem. In this research, we tested cross-lagged panel models (CLPMs) of the relation between the constructs (Finkel, 1995). Although we believe that the present analyses provide important information about the prospective effects between global and domain-specific self-esteem, it is important to acknowledge the ongoing debate about how prospective effects between constructs should be tested (e.g., Lucas, 2022; Lüdtke & Robitzsch, 2022; Orth, Clark, et al., 2021; Zyphur et al., 2020). For example, although Lüdtke and Robitzsch (2022) conclude, based on theoretical analyses and simulations, that the CLPM should be continued to be used, they suggest to control for lag-2 effects to strengthen the validity of conclusions (see also VanderWeele et al., 2020). Thus, to control for lag-2 effects, future research on the bottom-up and top-down models of self-esteem should use longitudinal data across at least three waves of assessment. Moreover, Hamaker et al. (2015) argue that prospective effects should be controlled for stable between-person differences in the constructs and therefore introduced the random-intercept CLPM (RI-CLPM). Again, when using the RI-CLPM, longitudinal data across at least three waves are needed. Consequently, in future research with multiwave longitudinal data, it would be interesting to compare the effects between global and domain-specific self-esteem across different models that have been suggested for testing prospective effects between constructs.

Conclusion

The present meta-analysis tested the bottom-up and top-down models of self-esteem, by comprehensively synthesizing the available longitudinal data. For most of the domains examined, the findings supported both bottom-up and top-down effects, suggesting that the relation between global and domain-specific self-esteem is best described by a reciprocal relation model. Moreover, the findings held across samples that differed with regard to age, gender, measure used for assessing domain-specific self-esteem, time lag, and publication year, which strengthens the generalizability of the results. Thus, the present research provides robust evidence for the long-assumed bottom-up and top-down processes between global and domain-specific self-esteem, as captured by Shavelson et al.'s (1976) hierarchical model of self-esteem.

References

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

Adachi, P., & Willoughby, T. (2015). Interpreting effect sizes when controlling for stability effects in longitudinal autoregressive models: Implications for psychological science. *European Journal of Developmental Psychology*, *12*(1), 116–128. https://doi.org/10.1080/17405629.2014.963549

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA publications and communications board task force report. *American Psychologist*, *73*(1), 3–25. https://doi.org/10.1037/amp0000191

*Asendorpf, J. B. (2006). Typeness of personality profiles: A continuous person-centred approach to personality data. *European Journal of Personality*, *20*(2), 83–106. https://doi.org/10.1002/per.575

*Barker, E. T., & Bornstein, M. H. (2010). Global self-esteem, appearance satisfaction, and self-reported dieting in early adolescence. *The Journal of Early Adolescence*, *30*(2), 205–224. https://doi.org/10.1177/0272431609332936

*Białecka-Pikul, M., Stępień-Nycz, M., Sikorska, I., Topolewska-Siedzik, E., & Cieciuch, J. (2019). Change and consistency of self-esteem in early and middle adolescence in the context of school transition. *Journal of Youth and Adolescence*, *48*(8), 1605–1618. https://doi.org/10.1007/s10964-019-01041-y

Blascovich, J., & Tomaka, J. (1991). Measures of self-esteem. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (Vol. 1, pp. 115–160). Academic Press. https://doi.org/10.1016/B978-0-12-590241-0.50008-3

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. https://doi.org/10.1002/9780470743386

*Boulton, M. J., Smith, P. K., & Cowie, H. (2010). Short-term longitudinal relationships between children's peer victimization/bullying experiences and self-perceptions: Evidence for reciprocity. *School Psychology International*, *31*(3), 296–311. https://doi.org/10.1177/0143034310362329

*Brendgen, M., Wanner, B., Vitaro, F., Bukowski, W. M., & Tremblay, R. E. (2007). Verbal abuse by the teacher during childhood and academic, behavioral, and emotional adjustment in young adulthood. *Journal of Educational Psychology*, *99*(1), 26–38. https://doi.org/10.1037/0022-0663.99.1.26

Brown, J. D. (1993). Self-esteem and self-evaluation: Feeling is believing. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 4, pp. 27–58). Lawrence Erlbaum.

Brown, J. D., & Dutton, K. A. (1995). The thrill of victory, the complexity of defeat: Self-esteem and people's emotional reactions to success and failure. *Journal of Personality and Social Psychology*, *68*(4), 712–722. https://doi.org/10.1037/0022-3514.68.4.712

Brown, J. D., Dutton, K. A., & Cook, K. E. (2001). From the top down: Self-esteem and self-evaluation. *Cognition and Emotion*, *15*(5), 615–631. https://doi.org/10.1080/02699930126063

Brown, J. D., & Marshall, M. A. (2006). The three faces of self-esteem. In M. H. Kernis (Ed.), *Self-esteem: Issues and answers* (pp. 4–9). Taylor & Francis.

Brummelman, E., & Sedikides, C. (2020). Raising children with high self-esteem (but not narcissism). *Child Development Perspectives*, *14*(2), 83–89. https://doi.org/10.1111/cdep.12362

Buer, L., Anderssen, N., Malterud, K., & Kvalem, I. L. (2017). Self-esteem among young bisexual women in Norway: A 13-year longitudinal study. *Journal of Bisexuality*, *17*(4), 487–507. https://doi.org/10.1080/15299716.2017.1394245

Byrne, B. M. (1996a). Academic self-concept: Its structure, measurement, and relation to academic achievement. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 287–316). Wiley.

Byrne, B. M. (1996b). *Measuring self-concept across the life span. Issues and instrumentation*. American Psychological Association. https://doi.org/10.1037/10197-000

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

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

*Chen, X., He, Y., & Li, D. (2004). Self-perceptions of social competence and self-worth in Chinese children: Relations with social and school performance. *Social Development*, *13*(4), 570–589. https://doi.org/10.1111/j.1467-9507.2004.00284.x

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple correlation/regression analysis for the social sciences*. Lawrence Erlbaum.

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

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, *19*(2), 300–315. https://doi.org/10.1037/a0033805

Cooper, H., Hedges, L. V., & Valentine, J. C. (2019). *The handbook of research synthesis and meta-analysis*. SAGE Publications. https://doi.org/10.7758/9781610448864

*Davidson, A. J., Gest, S. D., & Welsh, J. A. (2010). Relatedness with teachers and peers during early adolescence: An integrated variable-oriented and person-oriented approach. *Journal of School Psychology*, *48*(6), 483–510. https://doi.org/10.1016/j.jsp.2010.08.002

*Davison, K. K., Werder, J. L., Trost, S. G., Baker, B. L., & Birch, L. L. (2007). Why are early maturing girls less active? Links between pubertal development, psychological well-being, and physical activity among girls at ages 11 and 13. *Social Science & Medicine*, *64*(12), 2391–2404. https://doi.org/10.1016/j.socscimed.2007.02.033

Deković, M., van Aken, M. A. G., Reitz, E., van de Bongardt, D., Baams, L., & Doornwaard, S. M. (2018). *Project STARS. Studies on trajectories of adolescent relationships and sexuality*. Data Archiving and Networked Services (DANS). https://doi.org/10.17026/dans-z6b-t8ft

Diener, E., Northcott, R., Zyphur, M. J., & West, S. G. (2022). Beyond Experiments. *Perspectives on Psychological Science*, *17*(4), 1101–1119. https://doi.org/10.1177/17456916211037670

*Dohnt, H., & Tiggemann, M. (2006). The contribution of peer and media influences to the development of body satisfaction and self-esteem in young girls: A prospective study. *Developmental Psychology*, *42*(5), 929–936. https://doi.org/10.1037/0012-1649.42.5.929

*Donnellan, M. B., Trzesniewski, K. H., Conger, K. J., & Conger, R. D. (2007). A three-wave longitudinal study of self-evaluations during young adulthood. *Journal of Research in Personality*, *41*(2), 453–472. https://doi.org/10.1016/j.jrp.2006.06.004

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

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

*Dufner, M., Reitz, A. K., & Zander, L. (2015). Antecedents, consequences, and mechanisms: On the longitudinal interplay between academic self-enhancement and psychological adjustment. *Journal of Personality*, *83*(5), 511–522. https://doi.org/10.1111/jopy.12128

*Dundas, I., Thorsheim, T., Hjeltnes, A., & Binder, P. E. (2016). Mindfulness based stress reduction for academic evaluation anxiety: A naturalistic longitudinal study. *Journal of College Student Psychotherapy*, *30*(2), 114–131. https://doi.org/10.1080/87568225.2016.1140988

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *The BMJ*, *315*(7109), Article 629. https://doi.org/10.1136/bmj.315.7109.629

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

*Esnaola, I., Sesé, A., Antonio-Agirre, I., & Azpiazu, L. (2020). The development of multiple self-concept dimensions during adolescence. *Journal of Research on Adolescence*, *30*(Suppl. 1), 100–114. https://doi.org/10.1111/jora.12451

*Fenzel, L. M. (2000). Prospective study of changes in global self-worth and strain during the transition to middle school. *The Journal of Early Adolescence*, *20*(1), 93–116. https://doi.org/10.1177/0272431600020001005

Filosa, L., Alessandri, G., Robins, R. W., & Pastorelli, C. (2022). Self-esteem development during the transition to work: A 14-year longitudinal study from adolescence to young adulthood. *Journal of Personality*. Advance online publication. https://doi.org/10.1111/jopy.12713

Finkel, S. E. (1995). *Causal analysis with panel data*. SAGE Publications.

Fox, K. R., & Corbin, C. B. (1989). The physical self-perception profile: Development and preliminary validation. *Journal of Sport & Exercise Psychology*, *11*(4), 408–430. https://doi.org/10.1123/jsep.11.4.408

*Gallagher, A. M., Updegraff, K. A., Padilla, J., & McHale, S. M. (2018). Longitudinal associations between sibling relational aggression and adolescent adjustment. *Journal of Youth and Adolescence*, *47*(10), 2100–2113. https://doi.org/10.1007/s10964-018-0871-0

*Gest, S. D., Domitrovich, C. E., & Welsh, J. A. (2005). Peer academic reputation in elementary school: Associations with changes in self-concept

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

and academic skills. *Journal of Educational Psychology*, *97*(3), 337–346. https://doi.org/10.1037/0022-0663.97.3.337

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

Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, *58*(1), 80–92. https://doi.org/10.2307/1130293

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, *60*(1), 549–576. https://doi.org/10.1146/annurev.psych.58.110405.085530

Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The taboo against explicit causal inference in nonexperimental psychology. *Perspectives on Psychological Science*, *15*(5), 1243–1255. https://doi.org/10.1177/1745691620921521

Gruenenfelder-Steiger, A. E., Harris, M. A., & Fend, H. A. (2016). Subjective and objective peer approval evaluations and self-esteem development: A test of reciprocal, prospective, and long-term effects. *Developmental Psychology*, *52*(10), 1563–1577. https://doi.org/10.1037/dev0000147

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

Hardy, L., & Moriarty, T. (2006). Shaping self-concept: The elusive importance effect. *Journal of Personality*, *74*(2), 377–402. https://doi.org/10.1111/j.1467-6494.2005.00379.x

Harris, M. A., & Orth, U. (2020). The link between self-esteem and social relationships: A meta-analysis of longitudinal studies. *Journal of Personality and Social Psychology*, *119*(6), 1459–1477. https://doi.org/10.1037/pspp0000265

*Harris, M. A., Wetzel, E., Robins, R. W., Donnellan, M. B., & Trzesniewski, K. H. (2018). The development of global and domain self-esteem from ages 10 to 16 for Mexican-origin youth. *International Journal of Behavioral Development*, *42*(1), 4–16. https://doi.org/10.1177/0165025416679744

Harter, S. (1982). The perceived competence scale for children. *Child Development*, *53*(1), 87–97. https://doi.org/10.2307/1129640

Harter, S. (1998). The development of self-representations. In W. Damon & N. Eisenberg (Ed.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 553–617). Wiley.

Harter, S. (2003). The development of self-representations during childhood and adolescence. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 610–642). Guilford Press.

Harter, S. (2012a). *Self-perception profile for adolescents: Manual and questionnaires*. University of Denver.

Harter, S. (2012b). *Self-perception profile for children: Manual and questionnaires*. Univeristy of Denver, Department of Psychology.

Harter, S. (2015). *The construction of the self: Developmental and sociocultural foundations*. Guilford Press.

Harter, S., & Pike, R. (1984). The pictorial scale of perceived competence and social acceptance for young children. *Child Development*, *55*(6), 1969–1982. https://doi.org/10.2307/1129772

Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association*, *72*(358), 320–338. https://doi.org/10.1080/01621459.1977.10480998

Heatherton, T. F., & Polivy, J. (1991). Development and validation of a scale for measuring state self-esteem. *Journal of Personality and Social Psychology*, *60*(6), 895–910. https://doi.org/10.1037/0022-3514.60.6.895

*Helgeson, V. S., Escobar, O., Siminerio, L., & Becker, D. (2007). Unmitigated communion and health among adolescents with and without diabetes: The mediating role of eating disturbances. *Personality and Social Psychology Bulletin*, *33*(4), 519–536. https://doi.org/10.1177/0146167206296953

Jak, S., & Cheung, M. W.-L. (2020). Meta-analytic structural equation modeling with moderating effects on SEM parameters. *Psychological Methods*, *25*(4), 430–455. https://doi.org/10.1037/met0000245

Jak, S., Li, H., Kolbe, L., de Jonge, H., & Cheung, M. W. L. (2021). Meta-analytic structural equation modeling made easy: A tutorial and web application for one-stage MASEM. *Research Synthesis Methods*, *12*(5), 590–606. https://doi.org/10.1002/jrsm.1498

Kenny, D. A. (1979). *Correlation and causality*. Wiley.

*Kerr, D. C. R., Capaldi, D. M., Pears, K. C., & Owen, L. D. (2009). A prospective three generational study of fathers' constructive parenting: Influences from family of origin, adolescent adjustment, and offspring temperament. *Developmental Psychology*, *45*(5), 1257–1275. https://doi.org/10.1037/a0015863

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

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, *22*(17), 2693–2710. https://doi.org/10.1002/sim.1482

Kowalski, K. C., Crocker, P. R., Kowalski, N. P., Chad, K. E., & Humbert, M. L. (2003). Examining the physical self in adolescent girls over time: Further evidence against the hierarchical model. *Journal of Sport & Exercise Psychology*, *25*(1), 5–18. https://doi.org/10.1123/jsep.25.1.5

Krauss, S., & Orth, U. (2021). Work experiences and self-esteem development: A meta-analysis of longitudinal studies. *European Journal of Personality*. Advance online publication. https://doi.org/10.1177/08902070211027142

Langan, D., Higgins, J. P. T., Jackson, D., Bowden, J., Veroniki, A. A., Kontopantelis, E., Viechtbauer, W., & Simmonds, M. (2019). A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. *Research Synthesis Methods*, *10*(1), 83–98. https://doi.org/10.1002/jrsm.1316

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications.

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

Lucas, R. E. (2022). *It's time to abandon the cross-lagged panel model*. PsyArXiv. https://doi.org/10.31234/osf.io/pkec7

Lüdtke, O., & Robitzsch, A. (2022). A comparison of different approaches for estimating cross-lagged effects from a causal inference perspective. *Structural Equation Modeling*. Advance online publication. https://doi.org/10.1080/10705511.2022.2065278

Maes, M., Nelemans, S. A., Danneel, S., Fernández-Castilla, B., Van den Noortgate, W., Goossens, L., & Vanhalst, J. (2019). Loneliness and social anxiety across childhood and adolescence: Multilevel meta-analyses of cross-sectional and longitudinal associations. *Developmental Psychology*, *55*(7), 1548–1565. https://doi.org/10.1037/dev0000719

Marsh, H. W. (1986). Global self-esteem: Its relation to specific facets of self-concept and their importance. *Journal of Personality and Social Psychology*, *51*(6), 1224–1236. https://doi.org/10.1037/0022-3514.51.6.1224

Marsh, H. W. (1987). The hierarchical structure of self-concept and the application of hierarchical confirmatory factor analysis. *Journal of Educational Measurement*, *24*(1), 17–39. https://doi.org/10.1111/j.1745-3984.1987.tb00259.x

Marsh, H. W. (1990a). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review*, *2*(2), 77–172. https://doi.org/10.1007/BF01322177

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

Marsh, H. W. (1992a). *The Self Description Questionnaire (SDQ) I: A theoretical and empirical basis for the measurement of multiple*

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

dimensions of preadolescent self-concept: A test manual and research monograph. University of Western Sydney.

Marsh, H. W. (1992b). *Self Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept: An interim test manual and a research monograph*. University of Western Sydney.

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

Marsh, H. W. (1995). A Jamesian model of self-investment and self-esteem: Comment on Pelham (1995). *Journal of Personality and Social Psychology*, *69*(6), 1151–1160. https://doi.org/10.1037/0022-3514.69.6.1151

Marsh, H. W. (2008). The elusive importance effect: More failure for the Jamesian perspective on the importance of importance in shaping self-esteem. *Journal of Personality*, *76*(5), 1081–1122. https://doi.org/10.1111/j.1467-6494.2008.00514.x

Marsh, H. W., Barnes, J., Cairns, L., & Tidman, M. (1984). Self-Description Questionnaire: Age and sex effects in the structure and level of self-concept for preadolescent children. *Journal of Educational Psychology*, *76*(5), 940–956. https://doi.org/10.1037/0022-0663.76.5.940

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

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

*Marsh, H. W., Ellis, L. A., Parada, R. H., Richards, G., & Heubeck, B. G. (2005). A short version of the Self Description Questionnaire II: Operationalizing criteria for short-form evaluation with new applications of confirmatory factor analyses. *Psychological Assessment*, *17*(1), 81–102. https://doi.org/10.1037/1040-3590.17.1.81

Marsh, H. W., & O'Neill, R. (1984). Self description questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement*, *21*(2), 153–174. https://doi.org/10.1111/j.1745-3984.1984.tb00227.x

Marsh, H. W., & Redmayne, R. S. (1994). A multidimensional physical self-concept and its relations to multiple components of physical fitness. *Journal of Sport & Exercise Psychology*, *16*(1), 43–55. https://doi.org/10.1123/jsep.16.1.43

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist*, *20*(3), 107–123. https://doi.org/10.1207/s15326985ep2003_1

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

*McLeod, J. D., & Owens, T. J. (2004). Psychological well-being in the early life course: Variations by socioeconomic status, gender, and race/ethnicity. *Social Psychology Quarterly*, *67*(3), 257–278. https://doi.org/10.1177/019027250406700303

Messer, B., & Harter, S. (2012). *The self-perception profile for adults: Manual and questionnaires*. University of Denver.

*Mishna, F., Schwan, K. J., Lefebvre, R., Bhole, P., & Johnston, D. (2014). Students in distress: Unanticipated findings in a cyber bullying study. *Children and Youth Services Review*, *44*, 341–348. https://doi.org/10.1016/j.childyouth.2014.04.010

*Modecki, K. L., Blomfield Neira, C., & Barber, B. L. (2018). Finding what fits: Breadth of participation at the transition to high school mitigates declines in self-concept. *Developmental Psychology*, *54*(10), 1954–1970. https://doi.org/10.1037/dev0000570

*Morin, A. J., Maïano, C., Marsh, H. W., Janosz, M., & Nagengast, B. (2011). The longitudinal interplay of adolescents' self-esteem and body image: A conditional autoregressive latent trajectory analysis. *Multivariate Behavioral Research*, *46*(2), 157–201. https://doi.org/10.1080/00273171.2010.546731

*Narr, R. K., Allen, J. P., Tan, J. S., & Loeb, E. L. (2019). Close friendship strength and broader peer group desirability as differential predictors of adult mental health. *Child Development*, *90*(1), 298–313. https://doi.org/10.1111/cdev.12905

Neemann, J., & Harter, S. (2012). *Self-perception profile for college students: Manual and questionnaires*. University of Denver.

Niveau, N., New, B., & Beaudoin, M. (2021). Self-esteem interventions in adults–a systematic review and meta-analysis. *Journal of Research in Personality*, *94*(104131), 1–25. https://doi.org/10.1016/j.jrp.2021.104131

*Nogueira Avelar e Silva, R., van de Bongardt, D., Baams, L., & Raat, H. (2018). Bidirectional associations between adolescents' sexual behaviors and psychological well-being. *The Journal of Adolescent Health*, *62*(1), 63–71. https://doi.org/10.1016/j.jadohealth.2017.08.008

*Noordstar, J. J., van der Net, J., Jak, S., Helders, P. J., & Jongmans, M. J. (2016). Global self-esteem, perceived athletic competence, and physical activity in children: A longitudinal cohort study. *Psychology of Sport and Exercise*, *22*, 83–90. https://doi.org/10.1016/j.psychsport.2015.06.009

*O'Dea, J. A. (2009). Self perception score from zero to ten correlates well with standardized scales of adolescent self esteem, body dissatisfaction, eating disorders risk, depression, and anxiety. *International Journal of Adolescent Medicine and Health*, *21*(4), 509–517. https://doi.org/10.1515/IJAMH.2009.21.4.509

O'Mara, A. J., Marsh, H. W., Craven, R. G., & Debus, R. L. (2006). Do self-concept interventions make a difference? A synergistic blend of construct validation and meta-analysis. *Educational Psychologist*, *41*(3), 181–206. https://doi.org/10.1207/s15326985ep4103_4

Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2021). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*, *120*(4), 1013–1034. https://doi.org/10.1037/pspp0000358

Orth, U., Dapp, L. C., Erol, R. Y., Krauss, S., & Luciano, E. C. (2021). Development of domain-specific self-evaluations: A meta-analysis of longitudinal studies. *Journal of Personality and Social Psychology*, *120*(1), 145–172. https://doi.org/10.1037/pspp0000378

Orth, U., Erol, R. Y., & Luciano, E. C. (2018). Development of self-esteem from age 4 to 94 years: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *144*(10), 1045–1080. https://doi.org/10.1037/bul0000161

Orth, U., Meier, L. L., Bühler, J. L., Dapp, L. C., Krauss, S., Messerli, D., & Robins, R. W. (2022). Effect size guidelines for cross-lagged effects. *Psychological Methods*. Advance online publication. https://doi.org/10.1037/met0000499

Orth, U., & Robins, R. W. (2019). Development of self-esteem across the lifespan. In D. P. McAdams, R. L. Shiner, & J. L. Tackett (Eds.), *Handbook of personality development* (pp. 328–344). Guilford Press.

Orth, U., & Robins, R. W. (2022). Is high self-esteem beneficial? Revisiting a classic question. *American Psychologist*, *77*(1), 5–17. https://doi.org/10.1037/amp0000922

Orth, U., Robins, R. W., Widaman, K. F., & Conger, R. D. (2014). Is low self-esteem a risk factor for depression? Findings from a longitudinal study of Mexican-origin youth. *Developmental Psychology*, *50*(2), 622–633. https://doi.org/10.1037/a0033817

Pelham, B. W. (1995a). Further evidence for a Jamesian model of self-worth: Reply to Marsh (1995). *Journal of Personality and Social Psychology*, *69*(5), 1161–1165. https://doi.org/10.1037/0022-3514.69.6.1161

Pelham, B. W. (1995b). Self-investment and self-esteem: Evidence for a Jamesian model of self-worth. *Journal of Personality and Social Psychology*, *69*(6), 1141–1150. https://doi.org/10.1037/0022-3514.69.6.1141

Pelham, B. W., & Swann, W. B., Jr. (1989). From self-conceptions to self-worth: On the sources and structure of global self-esteem. *Journal of*

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.

Personality and Social Psychology, *57*(4), 672–680. https://doi.org/10.1037/0022-3514.57.4.672

*Preston, K. S., Gottfried, A. W., Oliver, P. H., Gottfried, A. E., Delany, D. E., & Ibrahim, S. M. (2016). Positive family relationships: Longitudinal network of relations. *Journal of Family Psychology*, *30*(7), 875–885. https://doi.org/10.1037/fam0000243

*Putnick, D. L., Hahn, C. S., Hendricks, C., & Bornstein, M. H. (2020). Developmental stability of scholastic, social, athletic, and physical appearance self-concepts from preschool to early adulthood. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *61*(1), 95–103. https://doi.org/10.1111/jcpp.13107

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (Vol. 2, pp. 295–316). SAGE Publications.

Reitz, A. K., Luhmann, M., Bleidorn, W., & Denissen, J. J. A. (2022). Unraveling the complex relationship between work transitions and self-esteem and life satisfaction. *Journal of Personality and Social Psychology*, *123*(3), 597–620. https://doi.org/10.1037/pspp0000423

Rentzsch, K., & Schröder-Abé, M. (2018). Stability and change in domain–specific self–esteem and global self–esteem. *European Journal of Personality*, *32*(4), 353–370. https://doi.org/10.1002/per.2167

Rentzsch, K., & Schröder-Abé, M. (2022). Top down or bottom up? Evidence from the longitudinal development of global and domain-specific self-esteem in adulthood. *Journal of Personality and Social Psychology*, *122*(4), 714–730. https://doi.org/10.1037/pspp0000393

Rentzsch, K., Wenzler, M. P., & Schütz, A. (2016). The structure of multidimensional self-esteem across age and gender. *Personality and Individual Differences*, *88*, 139–147. https://doi.org/10.1016/j.paid.2015.09.012

*Roeder, K. M., Cole, D. A., Sinclair, K. R., Dukewich, T. L., Preacher, K. J., Felton, J. W., Jacky, A., & Tilghman-Osborne, C. (2014). Sensitive periods for the effect of peer victimization on self-cognition: Moderation by age and gender. *Development and Psychopathology*, *26*(4pt1), 1035–1048. https://doi.org/10.1017/S0954579414000601

Rosenberg, M. (1979). *Conceiving the self*. Basic Books.

Rosenberg, M., Schooler, C., Schoenbach, C., & Rosenberg, F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review*, *60*(1), 141–156. https://doi.org/10.2307/2096350

*Rubie-Davies, C. M., & Hattie, J. A. (2012). The dangers of extreme positive responses in Likert scales administered to young children. *The International Journal of Educational and Psychological Assessment*, *11*(1), 75–89.

Scalas, L. F., Marsh, H. W., Nagengast, B., & Morin, A. J. (2013). Latent-variable approaches to the Jamesian model of importance-weighted averages. *Personality and Social Psychology Bulletin*, *39*(1), 100–114. https://doi.org/10.1177/0146167212465321

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

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

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

*Siffert, A., Schwarz, B., & Stutz, M. (2012). Marital conflict and early adolescents' self-evaluation: The role of parenting quality and early adolescents' appraisals. *Journal of Youth and Adolescence*, *41*(6), 749–763. https://doi.org/10.1007/s10964-011-9703-1

Sonstroem, R. J., & Morgan, W. P. (1989). Exercise and self-esteem: Rationale and model. *Medicine and Science in Sports and Exercise*, *21*(3), 329–337. https://doi.org/10.1249/00005768-198906000-00018

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

Steiger, A. E., Allemand, M., Robins, R. W., & Fend, H. A. (2014). Low and decreasing self-esteem during adolescence predict adult depression two decades later. *Journal of Personality and Social Psychology*, *106*(2), 325–338. https://doi.org/10.1037/a0035133

Swann, W. B., Jr., 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

Thompson, J. K., & Stice, E. (2001). Thin-ideal internalization: Mounting evidence for a new risk factor for body-image disturbance and eating pathology. *Current Directions in Psychological Science*, *10*(5), 181–183. https://doi.org/10.1111/1467-8721.00144

Trautwein, U. (2003). *Schule und Selbstwert*. Waxmann.

Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2006). Self-esteem, academic self-concept, and achievement: How the learning environment moderates the dynamics of self-concept. *Journal of Personality and Social Psychology*, *90*(2), 334–349. https://doi.org/10.1037/0022-3514.90.2.334

VanderWeele, T. J., Mathur, M. B., & Chen, Y. (2020). Outcome-wide longitudinal designs for causal inference: A new template for empirical studies. *Statistical Science*, *35*(3), 437–466. https://doi.org/10.1214/19-STS728

*Vanhalst, J., Luyckx, K., Scholte, R. H., Engels, R. C., & Goossens, L. (2013). Low self-esteem as a risk factor for loneliness in adolescence: Perceived—But not actual—Social acceptance as an underlying mechanism. *Journal of Abnormal Child Psychology*, *41*(7), 1067–1081. https://doi.org/10.1007/s10802-013-9751-y

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Viechtbauer, W. (2020). *metafor: Meta-analysis package for R* (Version 2.4-0). https://cran.r-project.org/package=metafor

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

Viechtbauer, W., López-López, J. A., Sánchez-Meca, J., & Marín-Martínez, F. (2015). A comparison of procedures to test for moderators in mixed-effects meta-regression models. *Psychological Methods*, *20*(3), 360–374. https://doi.org/10.1037/met0000023

Vohs, K. D., & Baumeister, R. F. (2012). *Self and identity (Vol. 1): Self-concept and self-esteem*. SAGE Publications.

*von Soest, T., Wichstrøm, L., & Kvalem, I. L. (2016). The development of global and domain-specific self-esteem from age 13 to 31. *Journal of Personality and Social Psychology*, *110*(4), 592–608. https://doi.org/10.1037/pspp0000060

Wagner, J., Lüdtke, O., Robitzsch, A., Göllner, R., & Trautwein, U. (2018). Self-esteem development in the school context: The roles of intrapersonal and interpersonal social predictors. *Journal of Personality*, *86*(3), 481–497. https://doi.org/10.1111/jopy.12330

*Wagnsson, S., Lindwall, M., & Gustafsson, H. (2014). Participation in organized sport and self-esteem across adolescence: The mediating role of perceived sport competence. *Journal of Sport & Exercise Psychology*, *36*(6), 584–594. https://doi.org/10.1123/jsep.2013-0137

Wang, Y., Chung, M. C., Wang, N., Yu, X., & Kenardy, J. (2021). Social support and posttraumatic stress disorder: A meta-analysis of longitudinal studies. *Clinical Psychology Review*, *85*, Article 101998. https://doi.org/10.1016/j.cpr.2021.101998

Webster, G. D., Howell, J. L., & Shepperd, J. A. (2022). Self-esteem in 60 seconds: The six-item state self-esteem scale (SSES-6). *Assessment*, *29*(2), 152–168. https://doi.org/10.1177/1073191120958059

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

Wichstrøm, L., & von Soest, T. (2016). Reciprocal relations between body satisfaction and self-esteem: A large 13-year prospective study of adolescents. *Journal of Adolescence*, *47*(1), 16–27. https://doi.org/10.1016/j.adolescence.2015.12.003

*Wilson, S., Schalet, B. D., Hicks, B. M., & Zucker, R. A. (2013). Identifying early childhood personality dimensions using the California Child Q-Set and prospective associations with behavioral and psychosocial development. *Journal of Research in Personality*, *47*(4), 339–350. https://doi.org/10.1016/j.jrp.2013.02.010

Zeigler-Hill, V. (2011). The connections between self-esteem and psychopathology. *Journal of Contemporary Psychotherapy*, *41*(3), 157–164. https://doi.org/10.1007/s10879-010-9167-8

Zyphur, M. J., Allison, P. D., Tay, L., Voelkle, M. C., Preacher, K. J., Zhang, Z., Hamaker, E. L., Shamsollahi, A., Pierides, D. C., Koval, P., & Diener, E. (2020). From data to causes I: Building a general cross-lagged panel model (GCLM). *Organizational Research Methods*, *23*(4), 651–687. https://doi.org/10.1177/1094428119847278

Received December 30, 2021
Revision received August 24, 2022
Accepted August 26, 2022 ■

This document is copyrighted by the American Psychological Association or one of its allied publishers.
Content may be shared at no cost, but any requests to reuse this content in part or whole must go through the American Psychological Association.