Arens, A. Katrin; Hasselhorn, Marcus

Age and gender differences in the relation between self-concept facets and self-esteem

Journal of Early Adolescence 34 (2014) 6, S. 760-791

Article

Journal of Early Adolescence
2014, Vol. 34(6) 760–791

DOI: 10.1177/0272431613503216
jea.sagepub.com

Age and Gender Differences in the Relation Between Self-Concept Facets and Self-Esteem

A. Katrin Arens[1] and Marcus Hasselhorn[1]

Abstract

This study tested whether the gender intensification hypothesis applies to relations between multiple domain-specific self-concept facets and self-esteem. This hypothesis predicts gender-stereotypic differences in these relations and assumes they intensify with age. Furthermore, knowledge about gender-related or age-related differences in self-concept–self-esteem relations might provide valuable knowledge for designing effective self-esteem enhancement interventions. We investigated grade and gender differences in the relations between domain-specific self-concept facets and self-esteem within a sample of 1958 German students in Grades 3 to 6. Results indicated no difference in the self-concept–self-esteem relations between the subsamples of third and fourth graders and fifth and sixth graders or between boys and girls. These relations also did not differ between boys and girls in the subsamples of third and fourth graders and fifth and sixth graders. These results suggest self-concept-self-esteem relations to be invariant across grade levels and gender and thus did not support the gender intensification hypothesis.

[1]German Institute for International Educational Research, Frankfurt, Germany

Corresponding Author:
A. Katrin Arens, Center for Research on Individual Development and Adaptive Education of Children at Risk, German Institute for International Educational Research, Schloßstraße 29, 60486 Frankfurt am Main, Germany.
Email: arens@dipf.de

Keywords
self-esteem, self-concept, age differences, gender differences

A high level of self-esteem in students has been found to be related to their psychological and physical well-being (e.g., Harter, 1990; Mann, Hosman, Schaalma, & de Vries, 2004; Trzesniewski et al., 2006). Numerous attempts have been made to investigate the determinants of self-esteem. Models of self-esteem determination often have assumed domain-specific self-concept facets to be sources of self-esteem and thus have examined the relations between various domain-specific self-concept facets and self-esteem (e.g., Harter, 1990; James, 1892). Some research has demonstrated gender-stereotypic patterns in the relations between domain-specific self-concept facets and self-esteem. For example, self-esteem was found to be more highly related to physical appearance self-concept for girls than for boys (Allgood-Merten, Lewinsohn, & Hops, 1990; Scalas & Marsh, 2008). According to the gender intensification hypothesis (Hill & Lynch, 1983), gender differences in the relations between domain-specific self-concept facets and self-esteem might intensify with age. In this study we examine grade level and gender differences in the relations between domain-specific self-concept facets and self-esteem to test the applicability of the gender intensification hypothesis to such relations. From a practical perspective, the investigation of age and gender differences in the relations between self-concept facets and self-esteem might provide knowledge required to design effective interventions for self-esteem enhancement that are appropriate for the age and gender of the target group.

Self-Concept and Self-Esteem

Self-concept has been defined as a multidimensional construct depicting self-perceptions in specific domains (Marsh, 1990; Marsh & Craven, 2006; Shavelson, Hubner, & Stanton, 1976). As such, students' self-concept can be divided into nonacademic and academic facets that both can be further differentiated into more specific self-concept domains. Nonacademic self-concept is supposed to consist of self-perceptions in physical, emotional, and social domains, and these self-concept facets can be yet further differentiated (e.g., physical self-concept comprises physical ability self-concept and physical appearance self-concept). Academic self-concept also comprises a domain-specific structure, as math and verbal self-concepts were found to form separate factors (Marsh, 1986, 1990). The strong domain specificity of academic self-concept (i.e., the separation between math and verbal facets) recently has been extended to the differentiation between competence and affect components. Hence, within each domain of academic self-concept (i.e., math and verbal self-concepts), a competence component depicting students' self-perceived competence and an affect component depicting students' motivational-affective reactions can be distinguished (Arens, Yeung, Craven, & Hasselhorn, 2011; Marsh, Craven, & Debus, 1999).

Domain-specific self-concept facets should be separated from self-esteem (also referred to as self-worth), which is commonly conceptualized as a domain-unspecific construct defining the global level of individuals' self-acceptance and self-respect (Harter, 1990; Marsh, 1990). A high level of self-esteem has been found to be associated with physiological and psychological well-being, whereas a low level has been found to be related to low motivation and behavioral misconduct (e.g., Harter, 1990; Trzesniewski et al., 2006). Furthermore, a low level of self-esteem is a risk factor for psychopathology including depression (e.g., Beck, Brown, Steer, Kuyken, & Grisham, 2001; Patton, 1991; Teri, 1982), anxiety (e.g., Beck et al., 2001; Ginsburg, La Greca, & Silverman, 1998), suicidal ideation (e.g., Harter, 1999), and eating disorders (e.g., Fisher, Pastore, Schneider, Pegler, & Napolitano, 1994; Stice, 2002). Given the desirable effects of a high level of self-esteem, it is worthwhile to investigate its determinants as such knowledge might facilitate the design of effective interventions for self-esteem enhancement. To gain insight into the determinants of self-esteem, researchers have investigated the relations between the domain-unspecific construct of self-esteem and domain-specific self-concept facets (Harter, 1999). A wide range of multidimensional self-concept facets have been found to be related to self-esteem whereby the strength of these relations have varied contingent on the specific self-concept facet in question. The strongest relations to self-esteem have been demonstrated for peer acceptance and physical appearance self-concepts (Frost & McKelvie, 2004; Harter, 1990; Marsh & Ayotte, 2003; Tiggemann, 2005; Wade, Thompson, Tashakkori, & Valente, 1989). Therefore, interventions targeting students' self-perceived peer acceptance might be an effective means for enhancing students' self-esteem. Accordingly, DeRosier (2004) conducted an experimental intervention study with children experiencing peer dislike, bullying, or social anxiety. Children of the experimental group demonstrated an increase in self-esteem after taking part in a social skills intervention program, whereas the control group demonstrated a decline in self-esteem. Due to the strong correlation between physical appearance and self-esteem found in previous studies (e.g., Harter, 1999), interventions addressing students' self-perceptions of physical appearance might also impact positively on self-esteem. Accordingly, Richardson and Paxton (2010) reported that adolescent girls had higher levels of self-esteem after a body image intervention. Thus, identifying the determinants of self-esteem by studying the relations between domain-specific self-concept facets and self-esteem might lead to important revelations about how to design effective interventions for self-esteem enhancement.

Relations between self-concept facets and self-esteem might change with students' age and might differ between boys and girls. Age and gender differences in domain-specific self-concept facets identified as determinants of self-esteem may lead to differential effectiveness of self-esteem interventions. If some domain-specific self-concept facets are closely related to self-esteem at some developmental stages but not at others, interventions aimed to improve self-esteem by addressing these self-concept facets will not be equally effective for students of all ages. Thus, knowledge about age-specific determinants of self-esteem as expressed by age-related changes in the relations between self-concept facets and self-esteem could be helpful in the designing of self-esteem interventions appropriate for various age groups. As the results of numerous studies showed a decline in students' self-esteem after transition to secondary school (Cantin & Boivin, 2004; Wigfield & Eccles, 1994; Wigfield, Eccles, MacIver, Reuman, & Midgley, 1991), self-esteem interventions seem to be needed particularly during this transition period. By studying the relations between domain-specific self-concept facets and self-esteem before and after transition to secondary school, researchers could gain insight into the determinants of students' self-esteem that might or might not change during transition. Such insight could help researchers and practitioners create self-esteem interventions that are particularly effective in counteracting the decrease in self-esteem during transition. Regarding gender differences, some self-concept facets might be related positively to girls' but not to boys' self-esteem. In this case, self-esteem enhancement programs targeted at domain-specific self-concept facets might not be equally effective for boys and girls. Hence, studying age and gender differences in the relations between self-concept facets and self-esteem could help researchers gain insight into age and gender differences in self-esteem determination and thus might help tailor self-esteem enhancement interventions to meet the specific age-related and gender-related needs of the target group.

Age Differences in Self-Esteem–Self-Concept Relations

Little research has been conducted on age differences in the relations between self-concept facets and self-esteem; however, these relations might vary with students' age. Due to cognitive-developmental advances, students might become increasingly able to establish abstract self-representations leading to a domain-unspecific sense of self (i.e., self-esteem; Harter, 1999), separated from domain-specific self-concept facets, resulting in weaker relations between domain-specific self-concept facets and self-esteem as students age. Age-related changes in the strength of the relations between domain-specific self-concept facets and self-esteem might also result from age-related shifts in the importance students attribute to specific life domains. The individual importance hypothesis (e.g., Hardy & Leone, 2008; James, 1892; Lindwall, Asci, Palmeira, Fox, & Hagger, 2011; Marsh, 2008) assumes that the relation between self-concept facets and self-esteem varies with the importance individuals allocate to specific self-concept domains. For example, physical ability self-concept is assumed to be strongly related to the self-esteem of individuals who deem sports of high subjective importance. In contrast, individuals who do not assign much importance to sports would not display a strong relation between physical ability self-concept and self-esteem. The individual importance hypothesis has originally assumed that individuals' subjective importance evaluations of specific life domains predict the salience of domain-specific self-concept facets for self-esteem. However, other approaches to the individual importance hypothesis take into consideration importance ratings resulting from individuals' group membership. Being a member of a social group means being influenced by the group's norms that might also comprise the level of importance the group members are expected to attribute to specific life domains. In other words, social groups often influence their members in terms of which life domains to deem important or unimportant. Consequently, self-concept facets related to life domains of high group importance are assumed to be strongly related to the self-esteem of all group members, whereas domains of little importance to the group might be only weakly related to group members' self-esteem. Accordingly, Hoge and McCarthy (1984) demonstrated that the weighting of specific self-concepts facets by group importance ratings was more effective than the weighting by individual importance ratings in predicting individuals' levels of self-esteem.

The relations between self-concept facets and self-esteem might change as students age, as the importance students allocate to different life domains may shift between childhood and adolescence. Shifts in importance ratings might result from changes in individuals' subjective importance hierarchy. In addition, shifts in importance ratings might originat from changes in students' group memberships, bringing along altered normative group importance ratings. For example, transition into adolescence with entry into the social group of adolescents might increase the importance of peer relations (e.g., Damon & Hart, 1988; Petersen, 1988). The increased importance of peer relations during adolescence might lead to stronger associations between self-esteem and peer self-concept. In addition, the relation between physical appearance self-concept and self-esteem might also become stronger during adolescence due to the growing importance of physical appearance at this age (e.g., Steinberg, 2005).

Gender Differences in Self-Esteem–Self-Concept Relations

In addition to age, gender might affect the relations between students' self-concept facets and self-esteem. Gender differences in the relations between self-concept facets and self-esteem might be due to gender differences in the importance attributed to specific life domains. Gender, a highly salient social group, is associated with gender role expectations that involve the importance of life domains. For example, congruent with the female gender stereotype, girls have been found to value physical appearance more than boys do, at least in Western societies (Jackson, Sullivan, & Hymes, 1987; Pliner, Chaiken, & Flett, 1990; Simmons & Blyth, 1987). Correspondingly, researchers have found a stronger connection between physical appearance self-concept and self-esteem for girls than for boys. For example, in the longitudinal study of Lord, Eccles, and McCarthy (1994), girls', but not boys', gains in self-esteem after transition to junior high school were associated with the level of self-perceived physical attractiveness prior to transition. Scalas and Marsh (2008) found that self-concept of physical appearance had a stronger indirect effect on women's self-esteem than on men's. The study of Allgood-Merten et al. (1990) also revealed stronger relations between self-perceived body image and self-esteem for girls than for boys. Similarly, in the qualitative study of Polce-Lynch, Myers, Kilmartin, Forssmann-Falck, and Kliewer (1998), adolescent girls reported more often than boys that their physical appearance contributed to their general feelings about themselves.

Females were also found to attribute higher importance to social relations and belongingness (e.g., Goodenow, 1993), which might be connected with the female gender stereotype of being nurturing, emotional and dependent (e.g., Del Boca & Ashmore, 1980). Correspondingly, some studies reported greater associations between social self-concept facets and self-esteem for girls than for boys. For example, Polce-Lynch et al. (1998) found that social experiences influenced students' feelings about themselves more for girls than for boys.

With regard to academic achievement, according to social norms, girls are expected to be more competent in verbal domains than boys, whereas boys are presumed to surpass girls in math achievement (e.g., Kiefer & Sekaquaptewa, 2007; Nosek, Banaji, & Greenwald, 2002). According to these social expectations, girls have been found to assign high importance to verbal academic skills (Jacobs, Hyatt, Osgood, Eccles, & Wigfield, 2002; Wigfield & Eccles, 1994). Corresponding to these gender differences in the importance ratings for verbal achievement, Byrne (1990; see also Byrne & Shavelson, 1987) demonstrated that girls' self-esteem was more strongly related to verbal (i.e., English) self-concept than to math self-concept.

As a high level of physical ability is often seen as a male characteristic, and boys have been found to have higher importance ratings for physical ability (e.g., Jacobs et al., 2002; Wigfield & Eccles, 1994), boys' self-esteem might have stronger relations to physical ability self-concept. Bowker (2006) provided corresponding empirical evidence as physical competence self-concept was a more significant predictor of boys' self-esteem than of girls' self-esteem.

Gender differences in the relations between domain-specific self-concept facets and self-esteem might be determined at least partly by gender-related differences in the importance attributed to specific life domains, which themselves might result from gender stereotypes and gender role expectations. However, not all findings regarding gender differences in the relation between self-concept facets and self-esteem correspond to the outlined gender differences in importance ratings. For example, although girls tend to deem social relationships as more important than boys do, Gecas and Schwalbe (1986) found adolescent boys' self-esteem to be more strongly related to self-perceptions of parent relations than girls' self-esteem. Similar findings were reported by Laible, Carlo, and Roesch (2004), who demonstrated that males' self-esteem was more strongly associated with parent attachment than was girls' self-esteem. Furthermore, Cairns, McWhirter, Duffy, and Barry (1990) reported that social self-concept is the best predictor of boys' but not girls' self-esteem. Overall, there is empirical evidence of gender differences in the relations between self-concept facets and self-esteem that are consistent with gender differences in importance ratings. However, as there are also findings of gender differences in the relations between self-concept facets and self-esteem that do not correspond to gender differences in importance ratings for self-concept facets, further research in this area seems to be necessary.

Age-Related Gender Differences in Self-Esteem–Self-Concept Relations

So far, we have discussed age-related and gender differences in the relations between domain-specific self-concept facets and self-esteem. More specifically, we have outlined how relations between self-concept facets and self-esteem might differ among students of different ages and between boys and girls. Integrating age and gender effects on the relations between self-concept facets and self-esteem can lead to the assumption that gender differences in these relations might vary with students' age. This assumption is evocative of the gender intensification hypothesis, which suggests increasing gender differences in psychological variables and processes during adolescence (Hill & Lynch, 1983; Maccoby, 1966). Concretely, the gender intensification hypothesis assumes that boys and girls increasingly internalize gender stereotypes and gender role expectations as they age. In concrete terms, girls are assumed to identify increasingly with female stereotypes and boys with male stereotypes. This process might lead to increasing gender differences in psychological outcomes resulting in girls' and boys' behaviors and attitudes becoming more akin to the corresponding gender stereotype.

Some studies have provided support for the gender intensification hypothesis by demonstrating an increasing gender gap in psychological processes and outcomes. For example, Galambos, Almeida, and Petersen (1990) found that gender differences in masculinity and sex roles increased between Grades 6 and 8. Furthermore, Roberts, Sarigiani, Petersen, and Newman (1990) demonstrated that the relationship between self-image and achievement decreased for girls but increased for boys between Grades 6 and 7. This finding was explained by boys' increasing adoption of achievement goals and girls' increasing adherence to social goals as expected by social norms. Regarding mean level differences in self-concept facets, the gender intensification hypotheses would assume boys to boost their self-concepts in male-related domains, such as math competence or physical ability, and girls to strengthen their self-concepts in female domains, such as verbal competence. However, empirical studies have shown that gender differences in the mean levels of various self-concept facets remain stable or even decrease throughout childhood and adolescence (Jacobs et al., 2002; Marsh, 1993; Marsh & Yeung, 1998a; Watt, 2004, 2008; Wigfield, Eccles, Yoon, Harold, Arbreton, & Blumenfeld, 1997). Hence, findings on the development of mean levels of self-concept facets are contrary to the assumptions of the gender intensification hypothesis.

To apply the gender intensification hypothesis to the relations between domain-specific self-concept facets and self-esteem would imply a gender-stereotypic pattern of relations that further intensifies with age. In concrete terms, the gender intensification hypothesis would predict that boys' self-esteem is more closely attached to male domains of self-concept (e.g., physical ability and math self-concepts) than girls' self-esteem, and that this gender gap in the strength of relations would increase with students' age. At the same time, girls are expected to display higher relations between self-esteem and self-concept facets corresponding to female stereotypes (e.g., physical appearance, social relations, and verbal self-concepts) than boys, whereby this gender difference in these relations might increase with age. Marsh (1993, 1994) tested age-related gender differences in the relations between various academic self-concept facets (general school, math, and verbal self-concepts) and self-esteem (referred to as general self-concept in his study) in students in Grades 7, 8, 9, and 10. The results of a series of invariance tests provided support for a gender-invariant model, as the relations between self-esteem and the measured academic self-concept facets did not differ between boys and girls in the different grade levels. However, as Marsh's (1993, 1994) studies considered only academic self-concepts and integrated only students in Grades 7 to 10, age-related gender differences in the relations between domain-specific self-concept facets and self-esteem cannot be ruled out when multiple nonacademic and academic self-concept facets and students of a younger age are investigated.

The Present Study

The aim of the present study is to explore grade level differences in the relations between domain-specific self-concept facets and self-esteem in German students when a wide range of academic and nonacademic self-concept facets is considered. We compared third and fourth graders with fifth and sixth graders with respect to the relations between various self-concept facets and self-esteem. By integrating German students before (third and fourth grade) and after (fifth and sixth grade) transition to secondary school, we were also able to examine changes in the relations between self-concept facets and self-esteem during transition to secondary school. In addition to grade level differences, we investigated gender differences in the relations between domain-specific self-concept facets and self-esteem and studied whether gender-specific patterns in these relations varied with age.

This study addresses theoretical and practical issues. By examining age, gender, and age-related gender differences in the relations between domain-specific self-concept facets and self-esteem, we provided a test for the applicability of the gender intensification hypothesis to relations between domain-specific self-concept facets and self-esteem when multiple academic and nonacademic self-concepts were taken into account. From a practical perspective, investigation into the relations between domain-specific self-concept facets and self-esteem might provide insight into the role of domain-specific self-concept facets as determinants of self-esteem. Researchers and practitioners are aware of the desirable effects of a high level of self-esteem on a wide range of behavioral and emotional outcomes (Harter, 1990; Mann et al., 2004; Trzesniewski et al., 2006) and have thus articulated a need for effective interventions to enhance students' self-esteem. Knowledge about the determinants of self-esteem could be used in the design of effective self-esteem enhancement interventions targeting domain-specific self-concepts. Further knowledge about whether and how the determinants of self-esteem vary contingent on students' age or gender might be helpful in the design of self-esteem enhancement programs to meet the needs of target groups what might further improve their effectiveness.

Method

Sample

The present study is based on a sample of 1,958 German students in Grades 3 to 6 (N = 989 [50.5%] boys and N = 969 [49.5%] girls) of 19 mixed-gender public schools located in the same geographical region in central Germany. To test whether the relations between boys' and girls' self-esteem and self-concept facets varied with age, the third and fourth graders (N = 617; N = 297 boys and N = 320 girls; age: M = 9.20, SD = 0.771) and the fifth and sixth graders (N = 1,341; N = 692 boys, N = 649 girls; age: M = 11.33, SD = 0.874) were combined in subsamples.

Students in Germany attend elementary school before they change to secondary school between Grades 4 and 5. Hence, the first subsample of third and fourth graders (N = 617) comprised elementary school students; the second subsample of fifth and sixth graders (N = 1,341) consisted of secondary school students. While in elementary school in Germany students are taught altogether regardless of their academic abilities, the transition to secondary school after Grade 4 coincides with an ability tracking procedure. After Grade 4, students are allocated to different secondary school tracks according to their accomplishments during the 4 years of elementary school. To obtain a sample that was sufficiently heterogeneous with regard to the cognitive abilities of the participating secondary school students, fifth and sixth graders from all secondary school tracks were included in this study: fifth grade: academic track: N = 240, intermediate track: N = 35, vocational track: N = 48, comprehensive track: N = 307; sixth grade: academic track: N = 196, intermediate track: N = 134, vocational track: N = 82, comprehensive track: N = 299.

The students in our sample came from 100 different classes (37 elementary school classes, 17 classes from the academic track of secondary school, 10 classes from the intermediate track of secondary school, 8 classes from the vocational track of secondary school, and 28 classes from the comprehensive track of secondary school). Only the students with parents' consent took part in the procedure. All of these students were included in the survey and no further selection criteria (e.g., based on students' ethnic background or cognitive abilities) were considered.

Instrumentation and Procedure

Students' domain-specific self-concept facets and self-esteem were measured using the short German version of the Self Description Questionnaire I (SDQ I-GS; Arens, Yeung, Craven, & Hasselhorn, 2013). The SDQ I-GS is considered an economical instrument for measuring the multidimensional self-concept of preadolescent children aged 8 to 12 and has demonstrated similar reliability and validity as a full-length German version of the SDQ I (Arens et al., 2013). The SDQ I-GS consists of 10 domain-specific scales for assessing different facets of students' nonacademic and academic self-concepts. With respect to nonacademic self-concept, students' self-concepts of physical appearance, physical ability, peer relations, and parent relations are assessed by separate scales. To measure academic self-concept, scales related to German language, math, and general school self-concepts are integrated in the instrument. While the German and math self-concept scales measure students' self-perceptions in these specific domains, general school self-concept refers to students' self-perceptions in all school subjects. The SDQ I-GS incorporates competence-related and affect-related items with respect to the different domains of academic self-concept considered (i.e., German, math, and general school self-concepts). The competence-related items ask for students' self-perceived competence with regard to German, math, and all subjects (e.g., "I am good at German/math/all school subjects") while the affect-related items depict students' affective and motivational reactions to German, math, and all school subjects (e.g., "I look forward to German/math/all school subjects"). The competence-related and affect-related items have been found to form separate factors representing separate competence and affect components within each domain-specific academic self-concept (i.e., German competence, German affect, math competence, math affect, general school competence, general school affect; Arens et al., 2011; Marsh et al., 1999). Besides the domain-specific academic and nonacademic self-concept scales, the SDQ I-GS consists of a separate scale for assessing students' self-esteem (also labeled as self-worth or general self-concept) in terms of self-respect and self-appreciation. Each scale of the SDQ I-GS instrument comprises three items. The items are worded as single-sentence statements. On a 5-point Likert-type scale, the students are asked to indicate whether each statement was true, mostly true, sometimes true/sometimes false, mostly false, or false.

The scales of the SDQ I-GS showed adequate reliability as indicated by the coefficient alpha estimates of internal consistency: physical appearance: $\alpha = .859$, physical ability: $\alpha = .876$, parent relations: $\alpha = .855$, peer relations: $\alpha = .801$, German competence: $\alpha = .875$, German affect: $\alpha = .889$, math competence: $\alpha = .905$, math affect: $\alpha = .934$, general school competence: $\alpha = .822$, general school affect: $\alpha = .879$, self-esteem: $\alpha = .748$. Given the recent critique of the coefficient alpha as an estimate of reliability (e.g., Sijtsma 2009), we also conducted estimates of scale reliability (ρ; also labeled as composite or instrument reliability) as an alternative index of reliability (Raykov, 2009). The estimates of scale reliability obtained for the SDQ I-GS were similar to the coefficient alpha reliability estimates: physical appearance: $\rho = .871$, physical ability: $\rho = .878$, parent relations: $\rho = .857$, peer relations: $\rho = .803$, German competence: $\rho = .876$, German affect: $\rho = .890$, math competence: $\rho = .905$; math affect: $\rho = .934$; general school competence: $\rho = .826$, general school affect: $\rho = .879$, self-esteem: $\rho = .750$.

The short German SDQ I was administered during students' regular classes by the first author or by a trained research assistant. The students were informed before they completed the self-concept measures that their data would be treated confidentially. The survey began by standardized instructions and example items for completing the SDQ I- GS. Afterward, each item was read aloud to circumvent confounding effects of reading ability. The students were asked to read along silently the items and to mark their chosen response on the 5-point Likert-type scale of the SDQ I-GS. Students without parents' consent left the classroom or were kept engaged with puzzles and paintings.

Statistical Analyses

The statistical analyses encompassed a stepwise procedure. A preliminary confirmatory factor analyses (CFA) model was used to test the factor structure of the SDQ I-GS. In an 11-factor model, each scale of the SDQ I-GS instrument was posited to form a separate factor. Hence, the 11-factor model (Model 1) consisted of 4 nonacademic self-concept factors (physical appearance, physical ability, parent relations, peer relations), 6 academic self-concept factors (German competence, German affect, math competence, math affect, general school competence, general school affect), and 1 self-esteem factor. The 11 self-concept factors were defined by the items of the corresponding self-concept scales only and no cross-loadings were permitted. Within the SDQ I-GS, some items for measuring the competence and affect components of the academic self-concept domains were worded in parallel across the domains of German language, math, and general school (e.g., "I am good at German/math/all school subjects"). Correlated uniquenesses between parallel worded items were permitted to take the shared measurement method of parallel worded items into account.

The research question of the present study regarding age, gender, and age-related gender differences in the relations between self-concept facets and self-esteem pertains to the concept of structural invariance. In general, tests of structural invariance examine whether certain aspects of the latent factors (i.e., factor variances, factor covariances, and latent means) are invariant across groups (Brown, 2006; Byrne, Shavelson, & Muthen, 1989). Tests of structural invariance are meaningful only if measurement invariance can be established a priori. This ensures that the observed (manifest) variables (i.e., the items of the SDQ I-GS scales in this study) measure the latent constructs (self-concept and self-esteem factors in this study) in the same way across groups (Brown, 2006). In other words, measurement invariance ascertains that the self-concept and self-esteem measures have the same underlying meaning across groups. For this reason, we started our analyses with tests of measurement invariance. To examine measurement invariance across grade levels, grade level (Grades 3 and 4 vs. Grades 5 and 6) was integrated as a grouping factor in Model 1 (i.e., in the 11-factor model). In a preliminary multigroup CFA Model (Model 4), an invariant factor pattern across the two subsamples of students' grade level was assumed (i.e., configural invariance; Meredith, 1993; Vandenberg & Lance, 1990). The same number of factors defined by the same set of items was stated for the subsample of third and fourth graders and for the subsample of fifth and sixth graders, but the factor loadings and item intercepts were freely estimated across groups. As a more restrictive model, the factor loadings were set equal across the two subgroups of grade levels in Model 5, but the item intercepts could vary across groups. The test of invariant factor loadings across groups (also known as weak measurement invariance, Meredith, 1993) is assumed to be the most important invariance test, as it provides the basis for all subsequent invariance tests (Brown, 2006). In Model 6 (model of strong measurement invariance; Meredith, 1993), the factor loadings and item intercepts were imposed to be the same across the two subgroups of grade level. As a model of strict measurement invariance with the factor loadings, item intercepts, and item uniquenesses constrained to be of equal size across grade levels, Model 7 completes the sequence of measurement invariance tests proposed by Meredith (1993).

After testing measurement invariance, we tested structural invariance. As the aim of the present study is to examine grade level and gender differences in the relations between domain-specific self-concept facets and self-esteem, the critical test is the examination of invariant relations between the domain-specific self-concept and self-esteem factors. Testing invariance of factor relations can be accomplished. By testing the invariance of factor covariances or the invariance of factor correlations (Marsh, 1994; Marsh & Hocevar, 1985). It is advised to start by examining the invariance of factor variances. In the case of invariant factor variances, tests of invariant factor covariances are justified, which are then equivalent to tests of invariant factor correlations (Brown, 2006; Marsh, 1994). Thus, in addition to invariant factor loadings, item intercepts, and item uniqueness (i.e., strict measurement invariance), Model 8 assumes invariant variances of SDQ I-GS factors. In Model 9, we tested whether the relations between the domain-specific self-concept factors and the self-esteem factor were invariant across grade levels by constraining the corresponding factor covariances to be the same across groups. Model 9 is thus a model of partial invariance, as it assumes only a subset of factor covariances (i.e., between the self-concept and self-esteem factors) to be equal across groups while another subset of factor covariances (i.e., the covariances among the self-concept factors) were freely estimated across groups.

All CFA models reported in this study were computed with the statistical package of Mplus, Version 7.0 (Muthén & Muthén, 1998-2012). The maximum likelihood estimator with robust standard errors (MLR) was chosen as the type of fitting function, as the categorical responses to the Likert-type scale of the SDQ I-GS instrument were treated as continuous variables. Missing values were estimated by the Full Information Maximum Likelihood (FIML) estimator implemented in Mplus. With only 0.72% items displaying missing values, the amount of missing data was very small in this study.

To evaluate the fit of the CFA models, several of the most common goodness-of-fit indices, such as the comparative fit index (CFI), the Tucker–Lewis index (TLI), the root mean square error of approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR) were used (Kline, 2005). With regard to the CFI and TLI, values above .90 are assumed to represent a good model fit (e.g., Kline, 2005). For the RMSEA, values below .05 are interpreted as a close fit, values between .05 and .08 as a reasonable fit, and values greater than .10 as a poor fit (Browne & Cudeck, 1993). For the SRMR, values below .10 are assumed to indicate a good model fit (Kline, 2005).

Examination of factorial invariance (i.e., measurement and structural invariance) involves comparing nested models that differ in their number of parameters restricted to be invariant across groups. Due to the sensitivity of the chi-square difference test to sample size, Cheung and Rensvold (2002) recommended using descriptive goodness-of-fit indices for comparing the fit of nested models. Factorial invariance (i.e., measurement and structural invariance) can be affirmed when the CFI does not drop more than .01, and the values of the RMSEA and SRMR do not increase more than .015 and .030, respectively, between models with less and more invariance constraints imposed. However, changes in the CFI should be taken as the main criterion for evaluating factorial invariance because changes in the RMSEA and SRMR may vary depending on the particular model parameters that are set invariant across groups (Chen, 2007).

After the sequence of models for testing grade level differences in the relations between self-concept facets and self-esteem, we turned to models for examining gender differences in these relations. For this purpose, we ran the same series of analyses as described above with gender included as a grouping variable. In essence, we tested the 11-factor model for boys (Model 10) and girls (Model 11) separately. Then, we conducted tests of measurement invariance including configural (Model 12), weak (Model 13), strong (Model 14), and strict (Model 15) measurement invariance. In the context of structural invariance testing, we used Model 16 to test the invariance of factor variances across gender. Finally, we used Model 17 to test gender differences in the relations between self-concept facets and self-esteem as in Model 17 the covariances between the self-concept and self-esteem factors were constrained to be equal across gender.

In a third set of analyses, we examined grade-related gender differences in the relations between self-concept facets and self-esteem. To this end, we conducted separate tests of measurement invariance across gender for students in Grades 3 and 4 (Models 20-23) and for students in Grades 5 and 6 (Models 28-31). These analyses were followed by tests of structural invariance including tests of invariant factor variances (Models 24 and 32) and tests of partial invariant factor covariances (Models 25 and 33).

Results

Grade Level Differences in Self-Concept–Self-Esteem Relations

The 11-factor model (i.e., Model 1 in Table 1) provided a good model fit for the total sample as indicated by all goodness-of-fit indices: χ^2 (426) = 1,051.894, CFI = .981, TLI = .976, RMSEA = .027, SRMR = .027. The self-concept factors were well defined as the standardized factor loadings were all substantial (ranging from .694 to .942, *Md* = .829). Therefore, we assumed the 11-factor model seemed to depict appropriately the multidimensional structure of students' self-perceptions (i.e., self-concept facets and self-esteem) measured by the SDQ I-GS.

The 11-factor model also demonstrated an adequate model fit when analyzing students in Grades 3 and 4 and students in Grades 5 and 6 separately

Table 1. Goodness-of-Fit Indices of the CFA Models.

Model	χ^2	*df*	CFI	TLI	RMSEA	SRMR	Model description
1	1,051.894	426	.981	.976	.027	.027	11-factor model for the total sample
Invariance tests across grade levels (3rd and 4th grades vs. 5th and 6th grades)							
2	654.662	426	.976	.970	.029	.033	11-factor model for 3rd and 4th grade students (*N* = 617)
3	911.691	426	.979	.974	.029	.029	11-factor model for 5th and 6th grade students (*N* = 1341)
4	1,565.143	852	.978	.973	.029	.030	Grade level as a grouping factor: no invariance constraints (configural invariance)
5	1,589.047	874	.978	.973	.029	.032	Grade level as a grouping factor: invariant factor loadings
6	1,649.386	896	.977	.973	.029	.032	Grade level as a grouping factor: invariant factor loadings, item intercepts
7	1,795.073	929	.973	.970	.031	.036	Grade level as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses
8	1,871.235	940	.971	.968	.032	.057	Grade level as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses, factor variances
9	1,901.166	950	.971	.967	.032	.059	Grade level as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses, factor variances, covariances between self-concept factors and the self-esteem factor
Invariance tests across gender							
10	803.133	426	.977	.972	.030	.032	11-factor model for boys (*N* = 989)
11	703.371	426	.983	.979	.026	.028	11-factor model for girls (*N* = 969)
12	1,507.837	852	.980	.975	.028	.030	Gender as a grouping factor: no invariance constraints (configural invariance)
13	1,534.369	874	.980	.976	.028	.031	Gender as a grouping factor: invariant factor loadings
14	1,637.075	896	.977	.973	.029	.032	Gender as a grouping factor: invariant factor loadings, item intercepts
15	1,720.973	929	.976	.973	.030	.036	Gender as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses
16	1,757.295	940	.975	.972	.030	.046	Gender as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses, factor variances

(continued)

Table 1. (continued)

Model	χ^2	*df*	CFI	TLI	RMSEA	SRMR	Model description
17	1,768.021	950	.975	.972	.030	.047	Gender as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses, factor variances, covariances between self-concept factors and the self-esteem factor
Grades 3 and 4 (*N* = 617)							
18	576.355	426	.970	.963	.034	.040	11-factor model for boys (*N* = 297)
19	571.499	426	.971	.964	.033	.042	11-factor model for girls (*N* = 320)
20	1,147.891	852	.970	.963	.034	.041	Gender as grouping factor: no invariance constraints (configural invariance)
21	1,178.182	874	.970	.963	.034	.045	Gender as a grouping factor: invariant factor loadings
22	1,222.237	896	.967	.962	.034	.047	Gender as a grouping factor: invariant factor loadings, item intercepts
23	1,302.971	929	.963	.958	.036	.053	Gender as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses
24	1,320.732	940	.962	.957	.036	.066	Gender as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses, factor variances
25	1,341.756	950	.961	.957	.037	.069	Gender as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses, factor variances, covariances between self-concept factors and the self-esteem factor
Grades 5 and 6 (*N* = 1,341)							
26	750.013	426	.973	.966	.033	.036	11-factor model for boys (*N* = 692)
27	652.606	426	.980	.976	.029	.030	11-factor model for girls (*N* = 649)
28	1,403.669	852	.976	.971	.031	.033	Gender as grouping factor: no invariance constraints (configural invariance)
29	1425.008	874	.976	.971	.031	.034	Gender as a grouping factor: invariant factor loadings
30	1,501.888	896	.974	.969	.032	.035	Gender as a grouping factor: invariant factor loadings, item intercepts
31	1,564.870	929	.973	.969	.032	.037	Gender as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses

(continued)

Table 1. (continued)

Model	χ^2	*df*	CFI	TLI	RMSEA	SRMR	Model description
32	1,591.998	940	.972	.969	.032	.047	Gender as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses, factor variances
33	1,600.245	950	.972	.969	.032	.049	Gender as a grouping factor: invariant factor loadings, item intercepts, item uniquenesses, factor variances, covariances between self-concept factors and the self-esteem factor
Secondary school tracks							
34	2,502.149	1,704	.968	.961	.037	.040	Secondary school track (academic track vs. other tracks) and gender as a joint grouping factor: no invariance constraints (configural invariance)
35	2,563.199	1,770	.968	.962	.037	.043	Secondary school track (academic track vs. other tracks) and gender as a joint grouping factor: invariant factor loadings
36	2,714.792	1,836	.965	.960	.038	.045	Secondary school track (academic track vs. other tracks) and gender as a joint grouping factor: invariant factor loadings, item intercepts
37	3,066.451	1,935	.955	.951	.042	.053	Secondary school track (academic track vs. other tracks) and gender as a joint grouping factor: invariant factor loadings, item intercepts, item uniquenesses
38	3,168.199	1,968	.952	.949	.043	.081	Secondary school track (academic track vs. other tracks) and gender as a joint grouping factor: invariant factor loadings, item intercepts, item uniquenesses, factor variances
39	3,204.445	1,998	.952	.949	.042	.084	Secondary school track (academic track vs. other tracks) and gender as a joint grouping factor: invariant factor loadings, item intercepts, item uniquenesses, factor variances, covariances between self-concept factors and the self-esteem factor

Note. All models were estimated by the maximum likelihood estimator with robust standard errors (MLR). CFI = Comparative Fit Index, TLI = Tucker–Lewis-Index, RMSEA = root mean square error of approximation, SRMR = standardized root mean squared residual.

(Models 2 and 3). Model 4 integrated the subsamples of students' grade levels (third and fourth vs. fifth and sixth grade) as a grouping factor and assumed the same factor pattern across these groups (i.e., configural invariance). This model also had a good model fit: χ^2 (852) = 1,565.143, CFI = .978, TLI = .973, RMSEA = .029, SRMR = .030. When assuming invariant factor loadings across groups (Model 5), the CFI value remained stable relative to Model 4 so that weak measurement invariance (i.e., invariant factor loadings) was supported. The assumption of invariant factor loadings and item intercepts (Model 6) also resulted in a nearly unchanged model fit, χ^2 (896) = 1,649.386, CFI = .977, TLI = .973, RMSEA = .029, SRMR = .032, indicating strong measurement invariance. Furthermore, the results supported strict measurement invariance, as the change in the CFI value (Δ = .004) was above Cheung and Rensvold's (2002) guideline for rejecting measurement invariance when additionally stating invariant item uniquenesses (Model 7). From the analyses conducted so far, measurement invariance of the SDQ I-GS across the two groups of grade levels (Grades 3 and 4 vs. Grades 5 and 6) can be seen as established allowing us to turn to tests of structural invariance. Above invariant factor loadings, item intercepts, and item uniquenesses, Model 8 posited invariant factor variances across grade levels. As the change in the CFI value by .002 (from .973 in Model 7 to .971 in Model 8) was above the guideline for rejecting invariance, the variances of the SDQ I-GS self-concept and self-esteem factors seemed to be invariant across grade level groups. Invariant factor variances allow testing the invariance of factor covariances (Brown, 2006; Marsh, 1994). Thus, to test whether the domain-specific self-concept factors were similarly related to the self-esteem factor in both grade level groups, we stated a model of partial invariance of factor covariances in Model 9. We constrained all the factor covariances addressing the relations between the various self-concept factors and the self-esteem factor (e.g., physical appearance self-concept and self-esteem, math competence self-concept and self-esteem, etc.) to be of equal size in the two subsamples of students' grade level. However, the covariances among the self-concept facets (e.g., physical appearance self-concept and math competence self-concept) could vary freely across groups. Compared with the fit of the less restrictive Model 8, the CFI and RMSEA values of Model 9 remained the same. This finding indicates similar relations between the various domain-specific self-concept facets and self-esteem for the groups of third and fourth graders and fifth and sixth graders. This result of invariant self-concept–self-esteem relations across grade levels does not correspond to the descriptive inspection of the correlations (Table 2), which demonstrates numerically more substantial coefficients for all but two (i.e., the relations of self-esteem to physical appearance and to parent relations self-concepts)

Table 2. Correlations of Domain-Specific Self-Concept Facets to Self-Esteem.

	Grades 3 and 4	Grades 5 and 6	Boys	Girls	Grades 3 and 4		Grades 5 and 6	
					Boys	Girls	Boys	Girls
	Model 4		Model 12		Model 20		Model 28	
Physical appearance	.601	.641	.645	.621	.642	.592	.654	.622
Physical ability	.450	.355	.423	.354	.551	.358	.373	.341
Parent relations	.428	.469	.471	.451	.495	.360	.458	.476
Peer relations	.809	.743	.769	.746	.767	.858	.774	.706
General school competence	.562	.514	.546	.525	.531	.583	.531	.491
General school affect	.538	.402	.462	.437	.520	.566	.428	.372
German competence	.464	.399	.482	.388	.521	.422	.454	.362
German affect	.380	.306	.405	.285	.445	.342	.375	.244
Math competence	.528	.307	.410	.347	.603	.471	.325	.284
Math affect	.449	.197	.279	.273	.468	.426	.187	.190

Note. Standardized solutions. All correlations are significant at $p < .001$.

relations between domain-specific self-concept facets and self-esteem for the third and fourth graders as compared with the fifth and sixth graders.

Gender Differences in Self-Concept–Self-Esteem Relations

We examined whether boys and girls differed in the relations between domain-specific self-concept facets and self-esteem. To test measurement invariance across gender, we stated a model of configural invariance (Model 12), a model of weak invariance (i.e., invariant factor loadings; Model 13), a model of strong measurement invariance (i.e., invariant factor loadings and item intercepts; Model 14), and a model of strict measurement (i.e., invariant factor loadings, item intercepts, and item uniquenesses, Model 15). As the goodness-of-fit indices remained nearly unchanged, and in particular the CFI did not drop more than .01 across these invariance models, measurement invariance across gender could be established as a prerequisite for testing structural invariance. Model 16, which also included invariant factor variances across gender, resulted in a CFI value that dropped by only .001 relative to the less restrictive Model 15. Given this evidence of invariant factor variances, Model 17 was used to test whether the different domain-specific self-concept facets were similarly related to self-esteem for boys and girls by assuming partial invariance of factor covariances. The CFI, TLI, and RMSEA values were the same for Models 16 and 17. Thus, the boys and girls of our sample did not seem to differ in the relations between the domain-specific self-concept factors and the self-esteem factor although inspection of Table 2 revealed consistently higher numerical relations for boys than for girls.

Age-Related Gender Differences in Self-Concept–Self-Esteem Relations

Results from the first two sets of analyses demonstrated that the third and fourth grade students displayed similar relations between domain-specific self-concept facets and self-esteem as did the fifth and sixth grade students. Furthermore, the relations between domain-specific self-concept facets and self-esteem seemed to be comparable between boys and girls. However, these findings do not preclude gender differences in the relations between self-concept facets and self-esteem that vary contingent on students' grade level. Thus, in the following, we examined gender differences in the relations between self-concept facets and self-esteem for third and fourth graders and then did the same for fifth and sixth graders.

Subsample of third and fourth grade students. The invariance models used with the subsample of third and fourth graders attested measurement invariance across gender. The descriptive fit indices did not change substantially between models assuming configural invariance (Model 20), weak measurement invariance (i.e., invariant factor loadings; Model 21), strong measurement invariance (i.e., invariant factor loadings and item intercepts, Model 22), and strict measurement invariance (i.e., invariant factor loadings, item intercepts, and item uniquenesses, Model 23). The results also supported invariance of factor variances (Model 24) given that the drop in the CFI ($\Delta = .001$ between Models 23 and 24) was negligible and above the guidelines of Cheung and Rensvold (2002) for rejecting invariance. Given the invariance of factor variances as a precondition, we tested invariance of the relations between domain-specific self-concept facets and self-esteem. Because the goodness-of-fit indices of the model of partial invariance of factor covariances (Model 25) demonstrated nearly unchanged descriptive fit indices (CFI $\Delta = .001$, RMSEA $\Delta = .001$, SRMR $\Delta = .003$), boys and girls attending Grades 3 and 4 were not found to display differential relations between self-concept facets and self-esteem. The descriptive results presented in Table 2 showed a tendency of higher relations between domain-specific self-concept facets and self-esteem for boys. Among the 10 assessed domain-specific self-concept facets, 7 were more highly related to boys' self-esteem than to girls' self-esteem (physical appearance, physical ability, parent relations, German competence, German affect, math competence, math affect).

Subsample of fifth and sixth grade students. For the subsample of fifth and sixth grade students, measurement invariance across gender could be demonstrated as in the sequence of configural invariance (Model 28), weak measurement invariance (Model 29), strong measurement invariance (Model 30), and strict measurement invariance (Model 31), the changes in the goodness-of-fit indices between more and less restrictive models were always above Cheung and Rensvold's (2002) guidelines for rejecting invariance. In addition, assuming invariant factor variances in Model 32 did not lead to a drop in model fit, which might lead one to argue against invariance. Because of the invariance of factor variances as a prerequisite, we could test invariance of relations between boys' and girls' multiple self-concept facets and self-esteem in Model 33. As the CFI value was the same for the less restrictive Model 32 (invariant factor loadings, item intercepts, item uniquenesses, and factor variances) and the more restrictive Model 33 (invariant factor loadings, item intercepts, item uniquenesses, factor variances, and partial factor covariances), the findings indicated gender invariant relations between the various self-concept facets and self-esteem for the students in Grades 5 and 6. However, similar to the findings for the third and fourth grade students, the descriptive results (Table 2) revealed that the boys tended to display stronger relations than the girls between domain-specific self-concept facets and self-esteem.

The fifth and sixth grade students were in different ability tracks of secondary school. To control for a potential bias due to the different tracks, we conducted supplementary analyses. We divided the sample of fifth and sixth graders into four groups: boys in the academic track ($N = 188$), boys in the other tracks (i.e., middle ability, low ability, and comprehensive tracks; $N = 504$), girls in the academic track ($N = 248$), and girls in the other tracks ($N = 401$). We integrated these subsamples of secondary school students as a grouping factor into a series of CFA models (Models 34-39 in Table 1). Compared with the model of configural invariance (Model 34), the CFI value remained the same when stating invariant factor loadings (Model 35). When stating invariant factor loadings and item intercepts (Model 36) or invariant factor loadings, item intercepts, and item uniquenesses (Model 37), the change in CFI value between the less and more restrictive models did not exceed the criterion of .01 for rejecting measurement invariance. Hence, measurement invariance across students in different secondary school tracks was established, allowing tests of structural invariance. As we also could demonstrate invariant factor variances (Model 38), Model 39 was used to test partial invariance of factor covariances regarding the relations between the different domain-specific self-concept facets and self-esteem. Due to the unchanged CFI value relative to Model 38, the four groups of boys and girls in different secondary school tracks did not differ in the strengths of the relations between domain-specific self-concept facets and self-esteem. Thus, the previous findings with regard to gender differences in the relations between self-concept facets and self-esteem in fifth and sixth graders were not biased according to their different ability tracks of the German secondary school system.

Discussion

The aim of the present study was to investigate grade level effects and gender differences in the relations between domain-specific self-concept facets and self-esteem and to investigate whether gender differences in these relations vary contingent on students' grade level. Hence, we tested the application of the gender intensification hypothesis (Hill & Lynch, 1983) to relations between multiple domain-specific self-concept facets and self-esteem. This study has practical implications, as the investigation into grade level and gender differences in the relations between students' self-concept facets and self-esteem might provide knowledge about the determinants of self-esteem that might vary with age and gender. This knowledge might be used to develop effective self-esteem intervention programs designed to meet the specific needs of various target groups.

Our conclusions are based on a series of invariance tests including models of measurement and structural invariance. The first set of analyses did not reveal any differences in the relations between self-concept facets and self-esteem between the subsamples of third and fourth grade and fifth and sixth grade students. Hence, irrespective of gender, German students in Grades 3 and 4 displayed relations between multidimensional self-concept facets and self-esteem similar to those of students in Grades 5 and 6. As the transition to secondary school takes place after Grade 5 in Germany, this finding implies that students before and after transition establish their self-esteem on the same domain-specific self-concept facets. The descriptive findings, however, revealed a tendency of stronger relations between self-concept facets and self-esteem for the subsample of third and fourth graders than for the subsample of fifth and sixth graders. Although this finding was not corroborated by the latent invariant tests, it might be the result of students' increasing ability to establish abstract self-representations (i.e., self-esteem) separated from domain-specific self-perceptions.

The second set of analyses did not evince differential relations between self-concept facets and self-esteem for boys and girls. Thus, irrespective of grade level, boys and girls displayed similar relations between the multidimensional domain-specific self-concept facets and self-esteem. This finding of gender-invariant relations counters the assumption that boys' self-esteem would reveal higher relations to male domains of self-concept (e.g., physical ability self-concept) and that girls' self-esteem would be more strongly related to female self-concept domains (e.g., physical appearance).

In the third set of analyses, we tested in particular the application of the gender intensification hypothesis to the relations between domain-specific self-concept facets and self-esteem. The gender intensification hypothesis assumes gender-stereotypic patterns of relations between self-concept facets and self-esteem that intensify with age. In neither of the groups (i.e., Grades 3 and 4 vs. Grades 5 and 6) did the boys and the girls reveal differential relations between self-concept facets and self-esteem. Although the invariance tests of this study indicated no gender differences in the relations between domain-specific self-concept facets and self-esteem, the descriptive analysis revealed higher correlations between domain-specific self-concepts and self-esteem for boys in the subsample of third and fourth graders and in the subsample of fifth and sixth graders. These findings might result from boys' greater reliance on domain-specific self-concept judgments for responding to domain-unspecific self-esteem items, or they might be a product of boys' less pronounced ability to establish abstract self-representations (i.e., self-esteem) detached from domain-specific self-perceptions. However, as these conjectures are speculative, further investigation is needed to unveil the reasons for the stronger correlations found in this study between self-esteem and domain-specific self-concept facets in boys than in girls.

In summary, the findings of this study did not provide support for the gender intensification hypothesis when applied to relations between domain-specific self-concept facets and self-esteem. In Grades 3 and 4, boys did not display stronger relations than girls between self-concept facets and self-esteem facets corresponding to the male stereotype, and these relations did not emerge or intensify in Grades 5 and 6. In parallel, in both subsamples of grade levels, girls' self-esteem was not found to be more highly related to self-concept domains corresponding to the female stereotype than were boys' self-esteem. Instead, for all students in this sample, and thus, irrespective of grade level and gender, self-concepts of physical appearance and peer relations displayed the strongest relations to students' self-esteem. This result corresponds to previous findings of strong relations of self-esteem to peer and physical appearance self-concepts (Harter, 1990; Wade et al., 1989). Hence, the salience of physical appearance and peer relations for self-esteem formation seems to be generalizable across gender and age. The strong connection between self-esteem and peer and physical appearance self-concepts, which seems to be established as early as in Grade 3, might prevent age-related and gender-stereotypic variations in the relations between self-concept facets and self-esteem. Further investigation into the emergence of the strong relation between self-esteem and physical appearance and peer self-concepts might be worthwhile. With respect to the practical implications of the results of this study, intervention programs for enhancing self-esteem should generally integrate strategies for fostering students' self-perceptions of physical appearance and peer acceptance.

Our findings support gender-invariant relations between multiple domain-specific self-concept facets and self-esteem for students in Grades 3 to 6. These findings are evocative of those presented by Marsh (1993, 1994) demonstrating gender-invariant relations between various academic self-concept facets (math, verbal, general school) and self-esteem for students in Grades 7 to 10. The results of this study seem to replicate as well as to extend Marsh's findings to multidimensional self-concepts including nonacademic and academic self-concept facets and to younger students in Grades 3 to 6. The results of the present study combined with those of Marsh (1993, 1994) provide evidence of gender-invariant relations between domain-specific self-concept facets and self-esteem in students from a wide age range countering the gender intensification hypothesis. This conclusion resembles results obtained from studies in which the gender intensification hypothesis was applied to the development of mean levels of self-concept. Those studies could not demonstrate that gender differences in the mean levels of self-concept facets increased with students' age; rather, they showed that such differences remained stable or even decreased with age (e.g., Jacobs et al., 2002; Marsh, 1993; Marsh & Yeung, 1998a; Watt, 2004, 2008; Wigfield et al., 1997). According to these findings along with the results of this study, the gender intensification hypothesis might not be applicable to students' multidimensional self-concept including the mean levels of its various facets and its relations to self-esteem.

By demonstrating gender-invariant relations between domain-specific self-concept facets and self-esteem, the findings of our study differ from studies demonstrating gender differences in the relations between domain-specific self-concept facets and self-esteem (e.g., Allgood-Merten et al., 1990; Bowker, 2006; Byrne, 1990; Gecas & Schwalbe, 1986). The divergence of findings might result from methodological shortcomings. In our study, as well in those of Marsh (1993, 1994) also showing gender-invariant self-concept–self-esteem relations, students' self-concepts and self-esteem were measured with the same instrument (i.e., the SDQ II in Marsh, 1993, 1994; the SDQ I-GS in this study). Thus, the invariant relations between self-concept facets and self-esteem found in this and Marsh's studies might be the result of method artifacts given due to the same instrument applied for assessing self-concept facets and self-esteem. For example, the chameleon effect (Marsh & Yeung, 1999) assumes that self-esteem ratings vary with the content of the other items of the survey the self-esteem items are embedded in. In this study, boys and girls of different ages might similarly have based their self-esteem ratings on the domain-specific self-concept ratings they were asked for in the SDQ I-GS, leading to the found gender-invariant relations between domain-specific self-concept facets and self-esteem. Hence, using different instruments to measure multidimensional self-concept facets and self-esteem would be useful for studying the issue of age and gender differences in the relations between them. Ideally, self-concept facets and self-esteem would be assessed at various points of time to preclude any transfer and method artifacts in the measurement of the different constructs.

In conclusion, our study provides some new and interesting findings regarding age and gender differences in the relations between domain-specific self-concept facets and self-esteem. With a large sample consisting of German students before and after transition to secondary school and by using the sophisticated methodological approach of latent invariance testing, the present study provides support for invariant relations between domain-specific self-concepts and self-esteem across age and gender, replicating previous findings (i.e., Marsh 1993, 1994). By considering a wide range of domain-specific self-concept facets, including nonacademic and academic facets, and by taking into account the recently found differentiability of academic self-concept facets into competence and affect components (Arens et al., 2011; Marsh et al., 1999), the present study furthers research on the relations between self-concept facets and self-esteem. Despite these strengths, our study has some shortcomings. Our data were cross-sectional; longitudinal studies would have provided better insight into the development of the relations between self-concept facets and self-esteem. Longitudinal studies could also investigate the causal relations between self-concept facets and self-esteem. Due to the cross-sectional nature of our data, we implicitly assumed reciprocal relations between self-concept facets and self-esteem. However, the reciprocal effects model is only one feasible model for the direction of the relation between self-concept facets and self-esteem. In addition to the reciprocal effects model, bottom-up models (self-concept determines self-esteem) and top-down models (self-esteem affects self-concept) have been discussed in the literature

(Marsh & Yeung, 1998b). Longitudinal data would have provided insight into the validity of these models and would have been useful in testing whether the causal flow between self-concept facets and self-esteem vary with students' age and with the specific self-concept facet in question. Furthermore, the sample of the study only covers Grades 3 to 6. Future research should involve students of a wider age range, as one might argue that age-related or gender-related changes in relations between domain-specific self-concept facets and self-esteem may occur at earlier or later stages of students' development. Moreover, future studies should take into consideration students' backgrounds (e.g., cognitive abilities, sociodemographic status) to control for variables that might impact on the strength of the relations between domain-specific self-concept facets and self-esteem. Given that the present study has been conducted in Germany where transition to secondary school takes place after Grade 5, studies including students from other countries and school systems with other times of transition would further contribute to the investigation of grade level, gender, and age-related gender differences in the relations between self-concept facets and self-esteem.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Allgood-Merten, B., Lewinsohn, P. M., & Hops, H. (1990). Sex differences and adolescent depression. *Journal of Abnormal Psychology*, *99*, 55-63.

Arens, A. K., Yeung, A. S., Craven, R. G., & Hasselhorn, M. (2011). The twofold multidimensionality of academic self-concept: Domain specificity and separation between competence and affect components. *Journal of Educational Psychology*, *103*, 970-981.

Arens, A. K., Yeung, A. S., Craven, R. G., & Hasselhorn, M. (2013). A short German version of the Self Description Questionnaire I: Theoretical and empirical comparability. *International Journal of Research & Method in Education*, *36*, 415-438.

Beck, A. T., Brown, G. K., Steer, R. A., Kuyken, W., & Grisham, J. (2001). Psychometric properties of the Beck Self-Esteem Scales. *Behaviour Research and Therapy*, *39*, 115-124.

Bowker, A. (2006). The relationship between sports participation and self-esteem during early adolescence. *Canadian Journal of Behavioural Science*, *38*, 214-229.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford.

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

Byrne, B. M. (1990). Investigating gender differences in adolescent self-concept: A look beneath the surface. *Applied Measurement in Education*, *3*, 255-274.

Byrne, B. M., & Shavelson, R. J. (1987). Adolescent self-concept: Testing the assumption of equivalent structure across gender. *American Educational Research Journal*, *24*, 365-385.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, *105*, 456-466.

Cairns, E., McWhirter, L., Duffy, U., & Barry, R. (1990). The stability of self-concept in late adolescence: Gender and situational effects. *Personality and Individual Differences*, *11*, 937-944.

Cantin, S., & Boivin, M. (2004). Change and stability in children's social network and self-perceptions during transition from elementary to junior high school. *International Journal of Behavioral Development*, *28*, 561-570.

Chen, F. F. (2007). Sensitivity of goodness of fit indices to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464-504.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*, 233-255.

Damon, W., & Hart, D. (1988). *Self-understanding in childhood and adolescence*. New York, NY: Cambridge University Press.

Del Boca, F. K., & Ashmore, R. D. (1980). Sex role stereotypes throughout the life cycle. In L. Wheeler (Ed.), *Review of personality and social psychology* (Vol. 1, pp. 163-192.). Beverly Hlls, CA: Sage.

DeRosier, M. E. (2004). Building relationships and combating bullying: Effectiveness of a school-based social skills group intervention. *Journal of Clinical & Adolescent Psychology*, *33*, 196-201.

Fisher, M., Pastore, P., Schneider, M., Pegler, C., & Napolitano, B. (1994). Eating attitudes in urban and suburban adolescents. *International Journal of Eating Disorders*, *16*, 67-74.

Frost, J., & McKelvie, S. (2004). Self-esteem and body satisfaction in male and female elementary school, high school, and university students. *Sex Roles*, *51*, 45-54.

Galambos, N. L., Almeida, D. M., & Petersen, A. C. (1990). Masculinity, femininity, and sex role attitudes in early adolescence: Exploring gender intensification. *Child Development*, *61*, 1905-1914.

Gecas, V., & Schwalbe, M. L. (1986). Parental behavior and adolescent self-esteem. *Journal of Marriage and the Family*, *48*, 37-46.

Ginsburg, G. S., La Greca, A. M., & Silverman, W. K. (1998). Social anxiety in children with anxiety disorders: Relations with social and emotional functioning. *Journal of Abnormal Child Psychology*, *26*, 175-185

Goodenow, C. (1993). Classroom belonging among early adolescent students: Relationships to motivation and achievement. *Journal of Early Adolescence*, *13*, 21-43.

Hardy, L., & Leone, C. (2008). Real evidence for the failure of the Jamesian perspective or more evidence in support of it? *Journal of Personality*, *76*, 1123-1136.

Harter, S. (1990). *Causes, correlates. and the functional role of global self-worth: A life-span perspective*. In R. J. Sternberg & J. Kolligian Jr. (Eds.), *Competence considered* (pp. 67-97). New Haven, CT: Yale University Press.

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

Hill, J. P., & Lynch, M. E. (1983). The intensification of gender-related role expectations during early adolescence. In J. Brooks-Gunn & A. C. Peterson (Eds.), *Girls at puberty* (pp. 201-228). New York, NY: Plenum Press.

Hoge, D. R., & McCarthy, J. D. (1984). Influence on individual and group identity salience in the global self-esteem of youth. *Journal of Personality and Social Psychology*, *47*, 403-414.

Jackson, L. A., Sullivan, L. A., & Hymes, L. A. (1987). Gender, gender role, and physical appearance. *The Journal of Psychology*, *121*, 51-56.

Jacobs, J. E., Hyatt, S., Osgood, W. D., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, *73*, 509-527.

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

Kiefer, A. K., & Sekaquaptewa, D. (2007). Implicit stereotypes, gender identification, and math-related outcomes. *Psychological Science*, *18*, 13-18.

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

Laible, D. J., Carlo, G., & Roesch, S. C. (2004). Pathways to self-esteem in late adolescence: The role of parent and peer attachment, empathy, and social behaviors. *Journal of Adolescence*, *27*, 703-716.

Lindwall, M., Asci, F. H., Palmeira, A., Fox, K. R., & Hagger, M. S. (2011). The importance of importance in the physical self: Support for the theoretically appealing but empirically elusive model of James. *Journal of Personality*, *79*, 303-334.

Lord, S. E., Eccles, J. S., & McCarthy, K. A. (1994). Surviving the junior high school transition. Family processes and self-perceptions as protective and risk factors. *Journal of Early Adolescence*, *14*, 162-199.

Maccoby, E. E. (1966). *The development of sex differences*. Stanford, CA: Stanford University Press.

Mann, M., Hosman, C. M. H., Schaalma, H. P., & de Vries, N. K. (2004). Self-esteem in a broad spectrum approach for mental health promotion. *Health Education Research*, *19*, 357-372.

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, *23*, 129-149.

Marsh, H. W. (1990). A multidimensional, hierarchical self-concept: Theoretical and empirical justification. *Educational Psychology Review*, *2*, 77-172.

Marsh, H. W. (1993). The multidimensional structure of academic self-concept: Invariance over gender and age. *American Educational Research Journal*, *30*, 841-860.

Marsh, H. W. (1994). Confirmatory factor analysis models of factorial invariance: A multifaceted approach. *Structural Equation Modeling*, *1*, 5-34.

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*, 1081-1121.

Marsh, H. W., & Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *Journal of Educational Psychology*, *95*, 687-706.

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*, 133-163.

Marsh, H. W., Craven, R. G., & Debus, R. (1999). Separation of competency and affect components of multiple dimensions of academic self-concept: A developmental perspective. *Merrill-Palmer Quarterly*, *45*, 567-601.

Marsh, H. W., & Hocevar, D. (1985). The application of confirmatory factor analyses to the study of self-concept: First and higher order factor structures and their invariance across age groups. *Psychological Bulletin*, *97*, 562-582.

Marsh, H. W., & Yeung, A. S. (1998a). Longitudinal structural equation models of academic self-concept and achievement. Gender differences in the development of math and English constructs. *American Educational Research Journal*, *35*, 705-738.

Marsh, H. W., & Yeung, A. S. (1998b). Top-down, bottom-up, and horizontal models: The direction of causality in multidimensional, hierarchical self-concept models. *Journal of Personality and Social Psychology*, *75*, 509-527.

Marsh, H. W., & Yeung, A. S. (1999). The lability of psychological ratings: The chameleon effect in global self-esteem. *Personality and Social Psychology Bulletin*, *25*, 49-64.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525-543.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide*. Seventh edition. Los Angeles, CA: Author.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal of Personality and Social Psychology*, *83*, 44-59.

Patton, W. (1991). Relationship between self-image and depression in adolescents. *Psychological Reports*, *68*, 687-688.

Petersen, A. C. (1988). Adolescent development. *Annual Review of Psychology*, *39*, 583-607.

Pliner, P., Chaiken, S., & Flett, G. L. (1990). Gender differences in concern with body weight and physical appearance over the life span. *Personality and Social Psychology Bulletin*, *16*, 263-273.

Polce-Lynch, M., Myers, B. J., Kilmartin, C. T., Forssmann-Falck, R., & Kliewer, W. (1998). Gender and age patterns in emotional expression, body image, and self-esteem: A qualitative analysis. *Sex Roles*, *38*, 1025-1048.

Raykov, T. (2009). Evaluation of scale reliability for unidimensional measures using latent variable modeling. *Measurement and Evaluation in Counseling and Development*, *42*, 223-323.

Richardson, S. M., & Paxton, S. J. (2010).An evaluation of a body image intervention based on risk factors for body dissatisfaction. A controlled study with adolescent girls. *International Journal of Eating Disorders*, *43*, 112-122.

Roberts, L. R., Sarigiani, P. A., Petersen, A. C., & Newman, J. L. (1990). Gender differences in the relationship between achievement and self-image during early adolescence. *Journal of Early Adolescence*, *10*, 159-175.

Scalas, L. F., & Marsh, H. W. (2008). A stronger latent-variable methodology to actual-ideal discrepancy. *European Journal of Personality*, *22*, 629-654.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Journal of Educational Research*, *46*, 407-441.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*, 107-120.

Simmons, R. G., & Blyth, D. A. (1987). *Moving into adolescence: The impact of pubertal change and school context*. Hawthorne, NJ: Aldine.

Steinberg, L. (2005). *Adolescence* (7th ed.). New York, NY: McGraw Hill.

Stice, E. (2002). Risk and maintenance factors for eating pathology: A meta-analytic review. *Psychological Bulletin*, *128*, 825-848.

Teri, L. (1982). Depression in adolescence: Its relationship to assertion and various aspects of self image. *Journal of Clinical Child Psychology*, *11*, 101-106.

Tiggemann, M. (2005). Body dissatisfaction and adolescent self-esteem: Prospective findings. *Body Image*, *2*, 129-135.

Trzesniewski, K. H., Donnellan, M. B., Moffitt, T. E., Robins, R. W., Poulton, R., & Caspi, A. (2006). Low self-esteem during adolescence predicts poor health, criminal behavior, and limited economic prospects during adulthood. *Developmental Psychology*, *42*, 381-390.

Vandenberg, R. J., & Lance, C. (1990). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations in organizational research. *Organizational Research Methods*, *3*, 4-70.

Wade, T. J. E., Thompson, V., Tashakkori, A., & Valente, E. (1989). A longitudinal analysis of sex by race differences in predictors of adolescent self-esteem. *Personality and Individual Differences*, *10*, 717-729.

Watt, H. M. G. (2004). Development of adolescents' self-perceptions, values, and task perceptions according to gender and domain in 7th through 11th grade Australian students. *Child Development*, *75*, 1556-1574.

Watt, H. M. G. (2008). A latent growth curve modeling approach using an accelerated longitudinal design: The ontogeny of boys' and girls' talent perceptions and intrinsic values through adolescence. *Educational Research and Evaluation*, *14*, 287-304.

Wigfield, A., Eccles, J., MacIver, D., Reuman, D. A., & Midgley, C. (1991). Transitions during early adolescence: Changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology*, *27*, 552-565.

Wigfield, A., & Eccles, J. S. (1994). Children's competence beliefs, achievement values, and general self-esteem: Change across elementary and middle school. *Journal of Early Adolescence*, *14*, 107-138.

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J., & Blumenfeld, P. C. (1997). Changes in children's competence beliefs and subjective task values across the elementary school years: A three-year study. *Journal of Educational Psychology*, *89*, 451-469.

Author Biographies

A. Katrin Arens is a postdoctoral fellow at the Center for Research on Individual Development and Adaptive Education of Children at Risk of the German Institute for International Educational Research. She completed her PhD in educational psychology in 2011. Her research interests are self-concept and motivation in education.

Marcus Hasselhorn is professor of psychology of education and human development and executive director of the German Institute for International Educational Research. He is also the scientific coordinator and speaker of the Center for Research on Individual Development and Adaptive Education of Children at Risk (IDeA) located in Frankfurt/Main, Germany. His research interests are the development of individual preconditions of successful learning, early education, and learning disabilities.