Self-Esteem, Academic Self-Concept, and Achievement: How the Learning Environment Moderates the Dynamics of Self-Concept

Ulrich Trautwein and Oliver Lüdtke
Max Planck Institute for Human Development

Olaf Köller
Humboldt University

Jürgen Baumert
Max Planck Institute for Human Development

The authors examine the directionality of effects between global self-esteem, domain-specific academic self-concepts, and academic achievement. Special emphasis is placed on learning environments as potential moderators of the direction of these effects. According to the meritocracy principle presented here, so-called bottom-up effects (i.e., self-esteem is influenced by academic self-concept) are more pronounced in meritocratic learning environments than in ego-protective learning environments. This hypothesis was examined using a three-wave cross-lagged panel design with a large sample of 7th graders from East and West Germany, a total of 5,648 students who were tested shortly after German reunification. Reciprocal effects were found between self-esteem, academic self-concept, and academic achievement. In conformance with the meritocracy principle, support for bottom-up effects was stronger in the meritocratic learning environment.

Keywords: self-esteem, self-concept, achievement, learning environments

A large body of research attests to desirable effects of strong beliefs in one’s own abilities and qualities on a variety of outcomes in settings such as health, education, and business (e.g., Bandura, 1997; Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005; Harter, 1998; Marsh, 1990). In recent years, researchers have increasingly begun to differentiate between global evaluations of the self, typically called "self-esteem," and evaluations of specific abilities or qualities, typically called “domain-specific self-concepts” (e.g., Bracken, 1996; Brown, 1993; Shavelson, Hubner, & Stanton, 1976; see below for more details). In the academic context, in particular, the differentiation between global self-esteem and domain-specific academic self-concepts has been shown to be important (e.g., M. Rosenberg, Schooler, Schoenbach, & F. Rosenberg, 1995; Valentine, DuBois, & Cooper, 2004).

Despite the obvious significance of self-concept in everyday life and research, however, the directionality of effects between self-esteem, domain-specific academic self-concept, and academic achievement remains hotly disputed (Baumeister, Campbell, Krueger, & Vohs, 2003; Brown, 1993; Dutton & Brown, 1997; Marsh & Yeung, 1998). Does high self-esteem (California Task Force to Promote Self-Esteem and Personal & Social Responsibility, 1990) or a high domain-specific academic self-concept (Marsh, Byrne, & Young, 1999; Marsh & Craven, in press) promote academic achievement? Is the global self-esteem of students influenced by their academic achievement and/or their domain-specific academic self-concepts (Harter, 1998; Shavelson et al., 1976)?

In this study, we tested the dynamics of self-concept in a large, representative sample of 7th grade students. In general, we expected to find reciprocal effects between self-esteem, self-concept, and achievement. However, we argue that the direction of the effects is also moderated by characteristics of the learning environment. More specifically, in accordance with the meritocracy principle, effects of domain-specific academic self-concepts on global self-esteem are expected to be more pronounced in meritocratic learning environments than in ego-protective learning environments.

The Directionality of Effects in Self-Concept: Theoretical Accounts

Empirical examination of the relationship between global self-esteem, academic self-concept, and academic achievement has been limited by two factors. First, researchers from different fields of study use different definitions of self-esteem, domain-specific academic self-concept, and achievement. Second, researchers with a background in personality and social psychology have typically focused on global self-esteem (e.g., Baumeister et al., 2003), whereas educational psychologists have typically concentrated on domain-specific self-concept (e.g., Marsh & Craven, in press). We therefore start by describing a hierarchical model of self-concept that covers all of the constructs used in this study. We then outline several theoretical approaches that illustrate the variety of theoretical positions on the directionality of effects between self-esteem,
domain-specific academic self-concept, and achievement that are advocated in the field.

The hierarchical model of self-concept by Shavelson et al., (1976) has attracted considerable research attention (e.g., Bracken, 1996; Byrne, 2002; Marsh & Shavelson, 1985; Marsh & Yeung, 1998). Because it also addresses academic achievement, this model is particularly useful in research that links self-concept to achievement (e.g., Skaalvik & Hagtvet, 1990). According to the model, self-concept is hierarchically structured, “with perceptions of behavior at the base moving to inferences about self in academic and nonacademic areas, and then to inferences about self in general” (Shavelson & Bolus, 1982, p. 3).

It has become common to refer to the general self-concept at the apex of the hierarchical self-concept model as global self-esteem (e.g., Marsh & Yeung, 1998; Rosenberg, 1965) or self-worth (Harter, 1998). In the present study, we use the term self-esteem to describe the global perception of the self as a person. People with a high level of self-esteem possess a positive overall view of the self. They appreciate who they are, consider themselves to have many good qualities, and would not want to change places with others (Baumeister et al., 2003; Rosenberg, 1965). People with low self-esteem, in contrast, have an unfavorable self-view. Self-esteem is very often measured by the Rosenberg (1965) Self-esteem Scale or a similar global scale from a self-concept inventory (for alternative conceptions, see Byrne, 1996; Marsh & Yeung, 1998). The items of the Rosenberg Self-esteem Scale (e.g., “All in all, I am inclined to feel that I am a failure”) and similar self-esteem scales tap global perceptions of the self as a person by exploring abstract, global self-evaluations. They do not target qualities in specific domains (e.g., academics, friendship). These domain-specific qualities and abilities are captured in the next levels of the hierarchical self-concept model by Shavelson et al. (1976).

In our study, in addition to self-esteem, we focused on domain-specific academic self-concepts. Domain-specific academic self-concepts reflect a person’s self-evaluation regarding a specific domain or ability in academic areas. Like global self-esteem, domain-specific self-concept is usually collected via self-report measures. Typical self-concept items are “I am quite good at mathematics” (mathematics self-concept) and “I have a poor vocabulary” (verbal self-concept). Domain-specific academic self-concepts are characterized by the more descriptive nature of the self-evaluations (compared with the more affective nature of the self-evaluations in global self-esteem). Domain-specific self-concept has also been called domain-specific self-esteem (e.g., Baumeister et al., 2003); however, in order to avoid any confusion with global self-esteem, we use the term academic self-concept throughout this article.

In contrast to the subjective nature of self-esteem and domain-specific academic self-concept, measures of academic achievement are intended to reflect the objective standing of a student. Teacher ratings of academic performance, school grades, and standardized achievement tests are typically used as measures of academic achievement.

What is the nature of the relationship between self-esteem, domain-specific academic self-concept, and academic achievement? In support of the hierarchical structure of self-concept, researchers have documented a closer relationship between self-esteem and academic self-concept than between self-esteem and academic achievement; likewise, academic achievement is more closely connected to academic self-concept than to self-esteem (see Byrne, 2002). More interesting, however, is the directionality of effects that these constructs have on one another over time. Does high self-esteem lead to a higher domain-specific self-concept over time if prior domain-specific self-concept is controlled? Does academic achievement influence the development of the respective domain-specific academic self-concept? Does mathematics self-concept have an impact on change in self-esteem? Drawing on hierarchical self-concept models in which self-esteem is at the apex of the hierarchy, the effects of self-esteem on academic self-concept and achievement and the effects of academic self-concept on achievement have been called top-down effects, whereas the effects of achievement on academic self-concept and self-esteem and the effects of academic self-concept on self-esteem have been dubbed bottom-up effects (see Marsh & Yeung, 1998). Yet the predominance of bottom-up versus top-down effects is still a matter of much debate.

Several psychological models have been developed to address this question. From a developmental perspective, Harter (1998) proposed a model that integrated the ideas of James (1892/1999) and Cooley (1902) to form a neo-Piagetian model of self-concept development. According to this model, global self-esteem results from one’s accomplishments in important areas and the feedback obtained from significant others (Harter & Whitesell, 1996), with the self-view becoming increasingly realistic from childhood to young adulthood. Although Harter acknowledged that global self-esteem can influence domain-specific self-evaluations (e.g., in the field of appearance), her model postulates the predominance of bottom-up effects (Harter, 1998). The bottom-up perspective is also emphasized in the model by Shavelson et al. (1976), according to which the self-concepts of a person “are formed through his experience with his environment...and are influenced especially by environmental reinforcements and significant others” (p. 411). Moreover, in the Shavelson et al. model, the impact of academic achievement on global self-esteem is mediated by domain-specific self-concept. Hence, Shavelson et al. emphasize the importance of bottom-up effects, though their model does not preclude the existence of top-down effects; in fact, later research by Shavelson and Bolus (1982) found support for such top-down effects.

Figure 1a is a schematic illustration of a strong bottom-up model for a two-wave example in the domain of mathematics. Academic achievement in mathematics at Time 1 (T1) has an impact on the corresponding domain-specific self-concept (mathematics self-concept) at Time 2 (T2). Similarly, mathematics self-concept (T1) impacts global self-concept at T2. These bottom-up effects are accompanied by horizontal effects that symbolize the stability of the three variables over time (also see Marsh & Yeung, 1998).

In contrast to the models by Harter (1998) and Shavelson et al., (1976), the theoretical and empirical work by Brown (e.g., Brown, 1986, 1993; Heimpel, Wood, Marshall, & Brown, 2002) places a strong emphasis on top-down effects. Brown described global self-esteem as a stable human characteristic that is formed early in life and has a major impact on human perception and action. He argued that the dominant flow of effects is from global self-esteem to more specific, domain-specific self-concepts. “Through transfer of affect processes and halo effects, positive feelings toward the self in general color people’s evaluations of their specific attributes” (Brown, 1993, p. 31). In a series of quasi-experiments
with university student samples, Brown and colleagues (1986, 1993; Brown & Dutton, 1995; Heimpel et al., 2002) have accumulated evidence for the influence of global self-esteem on cognitive, affective, and behavioral reactions to a variety of experimental stimuli. Figure 1b depicts Brown’s research hypotheses translated to a nonexperimental, longitudinal framework. In addition to horizontal effects, a positive effect of global self-esteem on domain-specific self-concept and academic achievement is expected.

A top-down perspective was also adopted by the California Task Force to Promote Self-esteem and Personal and Social Responsibility (1990). The members of this task force identified self-esteem as a major predictor of later academic accomplishment. Unfortunately, the members of the task force did not present an elaborated psychological model to describe the relevant processes that affect achievement, nor did they report adequate empirical support for their claim (see criticism by Baumeister et al., 2003). Moreover, although the emphasis of the report was on self-esteem, the distinction between self-esteem and domain-specific academic self-concepts was somewhat blurred in the description. Nevertheless, the task force report can be considered an example of a strong top-down perspective, as illustrated in Figure 1c.

More recent educational models (e.g., Eccles & Wigfield, 2002; Helmke & van Aken, 1995; Marsh et al., 1999) of the relationship between self-concept and academic achievement have typically focused on domain-specific academic self-concept rather than on global self-esteem. Most often, a reciprocal interrelationship is postulated, with both bottom-up and top-down effects (see Figure 1d). According to the expectancy-value theory (Eccles & Wigfield, 2002), for instance, self-concept is influenced by prior achievement-related experiences and feedback from relevant others, but it also has an impact on later achievement as a function of achievement-related choices and effort invested in learning. Similarly, Marsh and colleagues (e.g., Marsh & Craven, in press; Marsh et al., 1999) postulated positive effects of domain-specific self-concept on academic achievement and vice versa. With respect to global self-esteem, however, Marsh and colleagues (e.g., Marsh, 1993; Marsh & Hattie, 1996; Marsh & Yeung, 1998) have argued that global self-concepts have little bearing on specific self-concepts, that global self-esteem is not effective in explaining behavior when controlling for domain-specific self-concepts, and that domain-specific self-concepts have little predictive validity for self-esteem development. In the following section, we present empirical support for the various positions and report results from nonexperimental, large-scale longitudinal studies on the relationship between self-esteem, academic self-concept, and academic achievement.
Self-Esteem, Self-Concept, and Achievement: Empirical Studies

The focus of several studies has been the causal predominance of achievement, domain-specific self-concepts, and self-esteem. There is now considerable evidence for reciprocal effects between achievement and domain-specific academic self-concepts (see Marsh & Craven, in press; Valentine et al., 2004). Hence, domain-specific academic self-concepts are formed by academic achievement and influence later academic achievement. The effects found in longitudinal large-scale studies in which the same indicators are used in multiple waves of data collection tend to be rather small (typically \(0.05 < \beta < 0.15\)), but they do support the idea (Eccles & Wigfield, 2002; Marsh & Craven, in press) that academic self-concept and academic achievement have mutual positive effects.

When it comes to the relationship between global self-esteem and academic achievement, the pattern of empirical results is less definite. In a frequently cited review of literature, Baumeister et al. (2003) concluded that “the few studies suggesting any positive causal impact of self-esteem on academic achievement generally found only tiny effects. Some findings even point (again weakly) in the opposite direction, suggesting that high or artificially boosted self-esteem may detract from subsequent performance” (p. 13–14). Because of the dearth of longitudinal empirical studies, however, their conclusion was based on results from just five data sets.

Even fewer studies have been conducted on the dynamics between global self-esteem and domain-specific academic self-concepts from a nonexperimental, longitudinal perspective, and the results of these studies are also mixed. For instance, in a study reported by Skaalvik and Hagtvet (1990), academic self-concept had a small effect on self-esteem in the older cohort (students in Grades 6 and 7) but not in the younger cohort (students in Grades 3 and 4); no effects of self-esteem were found on later academic self-concept.

Studies that attest to the influence of global self-esteem on domain-specific self-concepts typically use quasi-experimental procedures (see Brown, 1993, for several examples). For example, the reactions of high-self-esteem students to a variety of treatments have been compared with those of low-self-esteem students. However, because most of these studies were conducted with college students, it is not clear whether these results would also hold in samples of children and adolescents.

To summarize, reciprocal effects have repeatedly been found between academic self-concept and achievement. There is less evidence for direct effects of self-esteem on achievement or vice versa. Finally, because few longitudinal studies have focused on the interrelationships between global self-esteem and domain-specific academic self-concepts, no final conclusions can be drawn about patterns of bottom-up and top-down effects between self-esteem and domain-specific academic self-concepts.

The Meritocracy Principle

The discussion so far has concentrated on the interrelationships between academic achievement, academic self-concept, and global self-esteem without taking the learning environment into account. However, the malleability of self-concepts when exposed to different contexts has been shown in several studies (see Wigfield, Eccles, & Pintrich, 1996, for a review). In the following text, we argue that different learning contexts might moderate the dynamics of self-concept, that is, the predominance of bottom-up or top-down effects between global self-esteem and domain-specific academic self-concepts. We introduce the meritocracy principle, which predicts that bottom-up effects (where self-esteem is influenced by academic self-concept) are more pronounced in meritocratic learning environments than in ego-protective learning environments.

Over the last 10–15 years, there has been an increasing awareness of the importance of classroom learning characteristics for students’ academic achievement, achievement-related behavior, and psychosocial well-being (e.g., Ames, 1992; Anderman, 2002; Pintrich, 2003). Although learning environments differ on several dimensions, the emphasis placed on effort, learning, competition and, more generally, achievement has been subject to particular scrutiny (see Pintrich, 2003). Little, Lopez, Oettingen, and Baltes (2001), for instance, reported a closer connection between academic achievement and academic self-concept in learning environments that emphasize accurate self-evaluations. The characteristics of the classroom learning environment also proved to be relevant in a study by Lüdtke, Köller, Marsh, and Trautwein (2005), who found mathematics self-concepts to be higher in classes in which teachers focused on individual students’ academic progress than in classes in which teachers stressed social comparisons. In terms of achievement-related behavior, it has been shown (Anderman & Midgley, 2004; Husemann, Trautwein, & Lüdtke, 2005) that cheating is more likely when students perceive their teachers to focus on performance (performance orientation) than when they emphasize student progress.

The meritocracy principle that we propose in this study postulates an effect of the learning environment on the relationship between academic self-concept and self-esteem. It holds that a student’s evaluation of his or her achievement in academic domains (academic self-concept) has a more pronounced impact on his or her global self-evaluation (self-esteem) in meritocratic learning environments than in less meritocratic learning environments. Meritocratic learning environments are characterized by two main features. First, they emphasize the importance of achievement, while at the same time valuing effort as a means to academic success. The use of unfair strategies to raise student achievement is considered unacceptable. Second, the “merits” of a student (his or her achievement and effort) are visible to classmates in meritocratic learning environments; social comparisons are used systematically and openly to show the meritocratic character of the learning environment and to heighten students’ motivation to excel. Hence, meritocratic environments place a strong emphasis on achievement and trigger social comparisons, but they also take account of the causes of success (e.g., Skinner, Chapman, & Baltes, 1988) by highlighting the key role of effort. We contrast meritocratic learning environments with what we call “ego-protective” learning environments. Here, academic success and failure are given less emphasis because social comparison plays a minor role and performance feedback tends to be given privately. Moreover, academic success and failure are seen as being multiply determined, with some of its causes falling outside a student’s realm of responsibility (e.g., difficulty of tasks), or being nonmeritocratic (e.g., ingratiating, luck).
What are the likely effects of meritocratic learning environments on the dynamics of self-concept? Several researchers have postulated the existence of bottom-up effects between academic self-concept and self-esteem (e.g., Harter, 1998; Shavelson et al., 1976). The meritocratic principle holds that the emphasis on achievement, effort, and social comparison in meritocratic learning environments is likely to increase these effects for several reasons. First, the emphasis on achievement, effort, and social comparison in meritocratic learning environments makes academic achievement a more important part of one’s global self-definition (Midgley, Feldlauer, & Eccles, 1989; Marsh, Köller, & Baumert, 2001). Second, the effects of academic self-concepts on self-esteem can be expected to increase because academic success or failure is heavily attributed to effort and, rather than attributing success or failure to external nonmeritocratic causes, teachers emphasize that students are responsible for their own achievement. Third, it can be speculated that the emphasis on social comparison increases students’ certainty in the accuracy of their academic self-concepts, whereas academic self-concepts in ego-protective learning environments might be more fragile. It can further be assumed that these more fragile academic self-concepts are receptive to the influences of high global self-esteem. In other words, there might be more of a balance between bottom-up and top-down effects between self-esteem and academic self-concepts in ego-protective environments.

Taken together, the characteristics of a meritocratic learning environment are likely to result in students’ self-concepts of academic ability having a relatively high impact on their formation of an overall, global self-evaluation. Conversely, the less emphasis is placed on effort and competition, and the more ingratiating is seen as a viable means to academic success, the less important bottom-up effects will be. The meritocratic principle is illustrated in Figure 1e. In the next section, we describe the East and West German school systems before reunification as examples of meritocratic (East Germany) and ego-protective (West Germany) learning environments in terms of students’ perceptions of the causes of high achievement in their classes.

The East and West German School Systems Before and After Reunification

For the present study, we used data collected in East and West Germany shortly after German reunification in 1990 to study the moderating effects of different contexts. Although based on the same cultural heritage, the school systems of the divided Germany reflected major societal differences (see Oettingen, Little, Lindenberger, & Baltes, 1994, for a thorough account of the differences in the school system).

In East Germany, there was no explicit streaming of students according to ability, and students attended the polytechnical high school (Polytechnische Oberschule) from Grade 1 to Grade 10. The West German system, on the other hand, tracked students of different ability, reflecting the liberal ideal that students should be offered a variety of school careers leading to different qualifications. Accordingly, there was early and selective assignment of students to different school tracks on the basis of 4 years’ performance at elementary school, resulting in more homogeneous learning environments in West Germany.

Oettingen et al. (1994) emphasized three major differences between the East and West German systems before reunification (see Marsh et al., 2001; Waterkamp, 1987, 1990). First, East German students were given detailed performance feedback in the form of grades from the first year of schooling, whereas West German students were given written reports without explicit grades for the first 2 years.

Second, in a related vein, it was an explicit educational goal of the East German system to foster accurate self-evaluations (Franz, 1987). The term “accurate self-evaluation” essentially meant that students were expected to agree with their teachers’ evaluation and was seen as a prerequisite for “harmoniously developed socialistic personalities” (Waterkamp, 1990, p. 263). Accordingly, teachers in East Germany emphasized social comparisons in their classes. Their feedback stressed the relative performances of individual students and was regularly given publicly in front of the entire class, as well as in public settings outside school. The West German system placed less emphasis on social comparison. In fact, performance feedback was more private, and students were not routinely informed of the grades attained by their classmates. Fostering accurate and highly differentiated self-evaluations was not an explicit goal of the West German school system. Rather, there was a focus on nurturing positive self-concepts.

Third, teachers in East Germany were required to implement a nationally valid curriculum reflecting the broader political program of the socialist government. Consequently, their teaching strategies and pace, as well as their educational goals and practices, were prescribed by the state. Achievement differences among students were attributed to differences in the effort that students invested in their work. Accordingly, students were expected to aspire to a high level of academic accomplishment by putting a great deal of effort into their learning, and teachers were expected to support student efforts to the best of their ability. In West Germany, teaching was subject to fewer regulations, and more emphasis was placed on students’ individual needs.

Empirical data indicate that academic achievement in East Germany compared favorably with student achievement in West Germany (Köller & Baumert, 2002). Yet soon after reunification, the East German states abandoned the unitary school system for political reasons, replacing it with the West German three-tier system. Hence, the East German school system was transformed, whereas the West German system changed very little (Wilde, 2002). Only a minority of East German teachers were replaced, however. Hence, there is reason to believe that the East and West German cultures of teaching and learning continued to differ in the period following reunification.

The Present Investigation

The interrelationship between global self-esteem, domain-specific academic self-concept, and academic achievement is of high theoretical and practical importance. Yet the exact nature of this relationship is far from being well understood (see reviews by Baumeister et al., 2003; Marsh & Craven, in press; Valentine et al., 2004). Although there is fairly sound support for a reciprocal relationship between academic achievement and domain-specific academic self-concepts, there has been little large-scale longitudinal research into the predominance of (a) global self-esteem versus academic achievement and (b) global self-esteem versus domain-specific academic self-concept.
In the present study, we had three main aims. First, we examined whether the different educational goals of the East and West German systems translated into different learning environments. More specifically, we examined the role that East and West German students attributed to effort and ingratiation as means of enhancing their achievement as indicators of meritocratic versus ego-protective learning environments. Second, and most important, we examined the dynamic interrelationship between domain-specific academic self-concepts and self-esteem, an important research area in which empirical evidence is still lacking (Baumeister et al., 2003; Marsh & Craven, in press; Valentine et al., 2004). At the core of our study was the question of whether characteristics of the learning environment have an impact on the dynamics of self-concept. More specifically, we tested the predictions of the meritocracy principle, according to which bottom-up effects would be stronger in meritocratic learning environments (i.e., the East German sample), where there is a focus on accurate self-evaluation, where social comparisons are emphasized, and where effort (rather than ingratiation) is seen as an important predictor of academic achievement, than in ego-protective learning environments (i.e., the West German sample).

Third, given the dearth of empirical data on the relationship between global self-esteem and achievement, we also examined the mutual effects of these two constructs from a longitudinal perspective. According to the review by Baumeister et al. (2003), we could have expected to find small top-down effects (if any); according to the findings of Harter (1998) and Shavelson et al. (1976), we could have expected to find more pronounced bottom-up effects. Although it is not at the core of the present study, we also examined the relationship between academic self-concepts and achievement. There is already sound research evidence for reciprocal effects between academic self-concept and achievement (Marsh & Craven, in press; Valentine et al., 2004). In accordance with this literature, we expected to find reciprocal effects between achievement and academic self-concept.

The present study drew on a strong database, which permitted our hypotheses to be studied in unusual depth. A large representative sample of 7th graders was surveyed three times over the course of 1 school year. In addition to school grades, standardized achievement measures were used as achievement indicators.

Method

Sample and Data

The empirical basis for this investigation was provided by the longitudinal study Learning Processes, Educational Careers, and Psychosocial Development in Adolescence and Young Adulthood (BIU) conducted at the Max Planck Institute for Human Development, Berlin. A total of 5,648 7th graders from 309 classes were tested on three occasions: at the beginning (T1), in the middle (T2), and at the end (T3) of the 1991–1992 school year. The sample consisted of 53.3% girls and 46.7% boys (mean age at T1 = 12.7 years). Student samples were drawn from three federal states (one in West and two in East Germany) at the beginning of the school year. To obtain representative samples of these three states, we stratified the sample by region and school type. Within each state and school type, random samples of schools were selected. Within each school, two 7th grade classes were randomly sampled. Overall, the sample comprised 50.8% West German and 49.2% East German students. Testing materials (standardized achievement tests and battery of psychological measures) were administered to intact classes by trained research assistants. T1 data collection coincided with the adoption of the West German secondary school system by the former East German states.

Instruments

Domain-specific self-concept. Mathematics and German self-concepts were measured by means of identical four-item scales (example item: “Nobody’s perfect, but I am just not good at mathematics [German]”). Students responded to each item in a four-point (agree–disagree) response format. Coefficient alphas were .81 (T1), .86 (T2), and .87 (T3) for mathematics self-concept and .70 (T1), .82 (T2), and .83 (T3) for German self-concept, respectively.

Self-esteem. Self-esteem was measured by a German short version (Jerusalem, 1984) of the Rosenberg Self-esteem Scale (Rosenberg, 1965). The short form consisted of four items (example item: “At times I think I am no good at all”). Again, a four-point scale (agree–disagree) was used. Coefficient alphas were .73 (T1), .80 (T2), and .81 (T3). An additional analysis with 210 7th–10th graders was performed to examine whether the German short version of the Rosenberg scale was representative of the complete German version (Ferring & Filipp, 1996). The latent correlation between the short form and the complete version equaled 1 (Trautwein, 2003).

Achievement tests, individual achievement scores, and school grades. The items constituting the mathematics achievement test were drawn from previous national and international studies, in particular the First and Second International Mathematics Study (FIMS, cf. Husén, 1967; SIMS, cf. Robitaille & Garden, 1989) developed by the International Association for the Evaluation of Educational Achievement and an investigation conducted at the Max Planck Institute for Human Development (cf. Baumert, Roeder, Sang, & Schnitz, 1986). The curricular validity of all items had been assessed beforehand by curriculum experts. The number of items was 30 at T1, 32 at T2, and 36 at T3. Various content areas were covered. The different test versions contained several identical items; that is, an anchor-item design (Hambleton & Swaminathan, 1989) was used. Internal consistency was above .80 at all points of measurement. Individual achievement scores were calculated in a vertical test-equating procedure on the basis of item response theory (see Birnbaum, 1968; Rasch, 1960). Using the computer program EQUATE 2.0 (Baker, 1994), the item and ability parameters of the second and third measurement points were rescaled according to the metric of the first point of measurement (see Köller, 1998, for more details of the scaling procedure).

In addition to mathematics achievement tests, school grades (mathematics, German) were collected from all students at T1 and T2. At T1, students reported their school grades at the end of the 6th grade; at T2, they reported the grades they had received on their Grade 7 midterm report card. In Germany, school grades are awarded on a six-point scale ranging from very poor to excellent. In the present study, the higher the grade value, the higher the achievement in the respective domain.

Learning environment. Students’ perceptions of meritocratic and non-meritocratic elements in their learning environments were collected via two scales at T1. These scales tapped students’ beliefs in effort and ingratiation as viable means of achieving academic success and were based in part on similar scales used by Nicholls, Cobb, Yackel, Wood, and Whealcy (1990). The effort scale consisted of six items (e.g., “You’re most likely to succeed at school if you work hard”); the ingratiation scale consisted of four items (e.g., “You’re most likely to succeed at school if you get in the teacher’s good books”). Internal consistency was satisfactory for both scales (effort: .81; ingratiation: .84).

Statistical Analyses

The main focus of the present study was threefold. First, we tested differences in the learning environment in East and West Germany. Our focus here is on mean level differences, but we also check for differences
in the correlations between our two measures (effort, ingratiation) across the two samples (East vs. West German students). Second, we examine the dynamics of self-esteem and domain-specific academic self-concept. Third, we inspect the pattern of mutual relationships between self-esteem (and academic self-concept) and achievement. In these latter sets of analyses, we focus on path coefficients relating the constructs.

Before our central analyses were carried out, however, we first examined the measurement equivalence of our main instruments across the two sociocultural contexts (East vs. West Germany). Although measurement equivalence might not be a necessary requisite of the analyses to be reported, an in-depth examination of the measurement equivalence across the two contexts might provide additional insights into any differences across the learning environments. Multiple-group mean and covariance structures (MACS) analyses (see Little, 1997) were thus performed. In confirmatory factor analysis studies with multiple groups, it is possible to test the invariance of any one, any set, or all parameter estimates across the multiple groups. To test for measurement invariance (see Jöreskog, 1979; Marsh, 1994), one analyzes a series of nested models, starting with the least restrictive model (no invariance constraints imposed) and ending with the most restrictive model (total invariance) in which all parameters are constrained to be the same across all groups. If the more restrictive models exhibit similar fit indices as the less restrictive models, the invariance of the constructs across the samples is supported.

In the present study, MACS analyses (Little, 1997) were used to establish the measurement equivalence (factor loadings and item intercepts) of all latent constructs (effort, ingratiation, self-esteem, mathematics and German self-concept) across East and West German samples. Furthermore, MACS analyses were used to examine our hypothesis that East German students would be more likely to attribute academic success to effort than their West German counterparts, whereas West German students would be more likely to ascribe success to ingratiation. Similarly, to test the assumed between-groups differences in the path coefficients relating achievement, self-concept, and self-esteem, we successively applied invariance constraints to the respective path coefficients.

We used the χ² statistic, the Tucker-Lewis Index (TLI; Bentler & Bonett, 1980), and the root-mean-square-error of approximation (RMSEA; Browne & Cudeck, 1993) to assess the fit of our models. Following the guidelines provided by Little (1997, pp. 58–59), measurement equivalence was evaluated using practical fit indices (i.e., TLI and RMSEA), whereas a statistical rationale (i.e., the χ² difference test) was applied when testing for systematic differences among the latent constructs (i.e., in the hypothesis-testing phase of our analyses). As is typical of studies conducted in the educational context, students were nested within classes and schools in the present study. Because students within the same class or school are on average more similar than students from different classes or schools, the standard errors resulting from routine standard procedures are typically downwardly biased. To correct for potential bias associated with class membership, we performed all MACS analyses with the computer program Mplus 3.1 (see Muthén & Muthén, 1998–2004), using the complex sample option. For the longitudinal analyses on the dynamics of self-concept, we followed Cronbach’s (1976) advice and standardized (M = 0.00, SD = 1.00) all variables within school classes to remove clustering effects. For the present study, class-mean centering was highly recommended because it helps to eliminate the potentially biasing effects of third variables at the class or school level, including the effects of different school types.

Missing values are a constant challenge in longitudinally oriented research projects because systematic drop-out threatens the generalizability of empirical results (Allison, 2001). In the present study, 5,648 students from 309 classes were sampled at T1. Because of organizational problems and the restructuring of the participating schools (e.g., relocation and reorganization of the East German schools), as well as small class sizes and refusal to participate, several classes could not be sampled at T2 (45 classes; 10.5% of the original participants) or T3 (71 classes; 19.8% of the original participants). A further 6.3% (T2) versus 7.7% (T3) of the original sample could not be resampled because of illnesses, relocation to new schools, or individual refusals. Altogether, data from 4,693 (83.1% of the original sample) versus 4,095 students (72.5% of original sample) were collected at T2 and T3, respectively (see Trautwein, 2003).

Recent research (Collins, Schafer, & Kam, 2001; see also Allison, 2001) suggests that powerful algorithms for substituting missing values, such as the expectation-maximization algorithm and multiple imputation, produce accurate estimations if data are missing at random. Even if data are not missing at random, using these algorithms results in less biased estimations than simpler procedures, such as listwise deletion. In the present study, the full information maximum likelihood estimator implemented in Mplus was used to deal with missing values. Mplus applies a model-based approach to missing data. Rather than missing values being imputed, missing data is recognized as such and all observed data values are used to estimate models with a full information maximum likelihood approach (see Allison, 2001).

Results

Applying MACS analyses (see Little, 1997), we first examined mean level differences between East and West German students’ perceptions of their learning environment. We then tested the measurement equivalence of our multiitem self-esteem and self-concept constructs across the two learning environments (East and West Germany) and studied the interrelationship between achievement, self-concept, and self-esteem in East and West Germany.

Perceived Characteristics of the Learning Environment

To test the measurement invariance of effort and ingratiation across the East and West German samples, we first estimated a model in which the measurement parameters in the two groups were simultaneously and freely estimated. The fit of this initial (unconstrained) model was good, χ²(68, N = 5,648) = 337.23, p < .001, TLI = 0.961, RMSEA = .040, suggesting that the hypothesized measurement model represented a good fit to the data in both groups. In the second model, we constrained all factor loadings to be equal in both samples. A test of the factorial invariance of the measurement model indicates whether the two constructs were understood similarly in both samples. The fit of this constrained model was still good, and the difference in fit was rather small, Δχ²(10) = 203.10, p < .001, ΔTLI = .025, ΔRMSEA = .009.¹ This result supports the view that the two constructs were comparable in both samples. In the third model, in addition to constraining the factor loadings to be equal, we imposed equality constraints on the intercepts of the measurement model. This fully constrained model showed a reasonable fit, χ²(86) = 726.68, p < .001, TLI = .927, RMSEA = .055, and the difference in fit with

¹ When one is using the complex option, Mplus provides a robust goodness-of-fit test. Because the difference between two robust chi-square variates does not have a chi-square distribution, we applied the procedure outlined in the technical appendix of the Mplus manual. This procedure computes a mean-adjusted robust chi-square difference test that uses a scaling factor c given by Mplus. Throughout the article, the reported significance levels are based on the corrected chi-square difference values.
the freely estimated model was acceptable, \( \Delta \chi^2(18) = 328.93, p < .001, \Delta \text{TLI} = .035, \Delta \text{RMSEA} = .015. \)

In the next step (the hypothesis-testing phase), we used the fully constrained model (factor loadings and intercepts invariant) to compare the latent means of the effort and ingratiation scales across East and West German students. Because all measurement parameters were held invariant across both groups, the means of the latent constructs can be meaningfully interpreted and compared. The latent means of the West German students were constrained to zero, whereas the latent means of East German students were freely estimated. Hence, the West German sample became the baseline sample against which the East German sample was compared. The latent means for the East German sample were \( M = 0.60 \) for effort and \( M = -0.57 \) for ingratiation. The difference in the means of the latent constructs between the two groups was highly significant, \( \Delta \chi^2(2) = 322.36, p < .001. \) Because we adjusted the standard deviation in both groups to 1, the difference between the means can be interpreted as standard effect size estimates. In other words, the East–West differences totaled more than half a standard deviation in both scales. East German students rated effort to be more important for achievement than their West German counterparts, whereas West German students rated ingratiation to be more important.\(^3\) Hence, in line with our hypothesis, the East German learning environment was experienced as substantially more meritocratic than the West German learning environment. The differences between East and West Germany amounted to more than half a standard deviation. It should be noted that differences of this size are rare in nonexperimental studies on learning environments.

We also examined the latent correlations among the two constructs (effort and ingratiation) in East and West Germany. Whereas there was a positive correlation of \( r = .17 \) (\( p < .001 \)) in the West German sample, a negative relationship of \( r = -.23 \) (\( p < .001 \)) was found in East Germany. The difference between the two latent correlations was again highly significant, \( \Delta \chi^2(1) = 86.70, p < .001. \) Hence, whereas West German students seemed to see effort and ingratiation as compatible means of achieving academic success, East German students differentiated more clearly between the two. Again, such pronounced differences in relationships across latent constructs are rare in nonexperimental field studies, and these differences in latent constructs further attest to the distinct nature of the learning environments in East and West Germany.

**The Dynamics of Achievement, Academic Self-Concept, and Self-Esteem**

We started our analyses of the dynamics between achievement, academic self-concept, and self-esteem by establishing measurement equivalence in the two samples (East and West German students). To constrain the factor loadings and item intercepts of the three latent variables (mathematics self-concept, German self-concept, global self-esteem) across the two samples (East and West German students) and the three measurement points (T1, T2, and T3), we computed a multiple-group structural equation model with six groups, in which every measurement point in each of the two contexts defined a different group. In line with Little (1997, see also Little et al., 2001), we again tested measurement equivalence by constraining the loadings and intercepts to equality across the samples but allowed the latent parameters to vary freely. Both the unconstrained model, \( \chi^2(306) = 1,927.00, \text{TLI} = .950, \text{RMSEA} = .047, \) and the fully constrained model (factors loadings and item intercepts constrained to equality), \( \chi^2(396) = 3,143.73, \text{TLI} = .935, \text{RMSEA} = .054, \) exhibited a good fit. According to the guidelines provided by Little (1997), the difference in fit between the unconstrained and the constrained models was reasonably small, \( \Delta \chi^2(90) = 928.47, \Delta \text{TLI} = .015, \Delta \text{RMSEA} = .007. \) Hence, the analyses reported thus far support the equivalence of measurements in the two groups and at each time point.

We next tested for context effects as moderators of the self-concept dynamic (the hypothesis-testing phase). A set of cross-lagged panel analyses was specified, with particular emphasis being placed on the cross-lagged effects that would constitute bottom-up effects (achievement \( \rightarrow \) academic self-concept; achievement \( \rightarrow \) self-esteem; academic self-concept \( \rightarrow \) self-esteem) or top-down effects (self-esteem \( \rightarrow \) academic self-concept; self-esteem \( \rightarrow \) achievement; academic-self-concept \( \rightarrow \) achievement). By placing equality constraints on these bottom-up and top-down effects, we were able to test for East–West differences in the dynamics of self-concept.

Examining the dynamics between achievement, self-concept, and self-esteem, we specified separate sets of models for mathematics and German, with achievement, self-concept, and self-esteem measured at an earlier time point predicting the same achievement indicators (standardized achievement tests, school grades). In all two-group models reported, factor loadings were held equal across the two groups. Residual covariances between identical items over the three time points were freely estimated (see Marsh & Hau, 1996).

Because preliminary analyses showed positive correlations among the three constructs (achievement, academic self-concept, and self-esteem), multicollinearity might obscure the pattern of results. For instance, both global self-esteem and mathematics self-concept might have significant effects on mathematics

\(^2\) Although our tests of the overall model supported the assumption of measurement invariance, we performed additional analyses to test the misfit of individual model parameters (“local misfit”). These analyses identified the fit of one item from the effort scale to be less satisfactory: “You’re most likely to succeed at school if you prepare well for tests and examinations.” The difference between the mean scores of East and West German students on the 4-point scale was very pronounced for this item (\( M = 3.69 \) vs. \( M = 3.27 \)), and there was little variance in the East German sample (\( SD = 0.56 \) vs. \( SD = 0.88 \)). It is thus possible that the misfit for this item can be attributed to a ceiling effect among the East German students. We also tested for measurement invariance without this item. The deterioration in the goodness of fit between the unconstrained model, \( \chi^2(52) = 278.76, p < .001, \text{TLI} = .959, \text{RMSEA} = .042, \) and the model in which factor loadings and intercepts were constrained to be equal, \( \chi^2(68) = 505.20, p < .001, \text{TLI} = .939, \text{RMSEA} = .051, \) was comparatively small, \( \Delta \chi^2(16) = 182.05, p < .001, \Delta \text{TLI} = .020, \Delta \text{RMSEA} = .009. \)

\(^3\) There was little change in the differences in the latent means between East and West German students when the models were computed again without the item with the greatest misfit (latent means for East German sample: effort, \( M = 0.55 \); ingratiation, \( M = -.057 \)). The mean differences between East and West German students were again significant, \( \Delta \chi^2(2) = 266.12, p < .001. \)
achievement when considered separately, though the effects of neither are statistically significant when both constructs are considered simultaneously (i.e., the unique effects of neither are significant when the effects of each are controlled for the effects of the other). Hence, we started with a model in which self-esteem, mathematics self-concept, and mathematics achievement were included (see Figure 2), but we subsequently computed a series of supplemental models to evaluate the directionality of effects among various pairs of constructs (see Table 1 and Figures 3a–3h).

The two-group model (West vs. East German students), including mathematics achievement (standardized test scores), mathematics self-concept, and self-esteem, exhibited a good fit, $\chi^2(588) = 1,402.59$, TLI = .978, RMSEA = .022 (see Model 1a in Table 1). Figure 2 reports the standardized regression coefficients for this three-wave cross-lagged panel analysis, separately for the East and West German students. As expected, the largest coefficients were found for the stability coefficients: T1 constructs significantly and meaningfully predicted the corresponding T2 constructs, and T2 constructs predicted the corresponding T3 constructs.

Most important in the present context, however, we also found several significant bottom-up and top-down effects between self-esteem and mathematics self-concept. These effects were fully in line with the meritocracy principle. T1 mathematics self-concept predicted global self-esteem in the East German sample only: The effect of T1 mathematics self-concept on T2 self-esteem was $\beta = .02$ (ns) in the West German sample, compared with $\beta = .16$ ($p < .001$) in the East German sample. Similarly, the effect of T2 mathematics self-concept on T3 self-esteem was significant in the East German sample only (West: $\beta = -.01$, ns; East: $\beta = .14$, $p < .001$). Conversely, significant top-down effects of self-esteem on mathematics self-concept were more consistent in the West than in the East German sample (T1/T2 for West: $\beta = .09$, $p < .001$; East: $\beta = .04$, ns. T2/T3 for West: $\beta = .10$, $p < .01$; East: $\beta = .07$, $p < .01$).

To test these East–West differences statistically, we reran the structural equation model with additional equality constraints placed on the respective bottom-up and top-down paths. As documented in Table 1, when bottom-up effects were constrained to be equal across the two samples (see Model 1b), the goodness of fit was significantly lower than in Model 1, $\Delta \chi^2 (6) = 37.53$, $p < .001$. No such effect was found for the top-down effects (see Model 1c). Hence, taken together, the results attest to moderating effects of the two learning environments on the dynamics between self-esteem and academic self-concept.

In Model 1a, we also found evidence for a reciprocal relationship between mathematics achievement and mathematics self-concept, with prior self-concept significantly predicting later achievement and prior achievement significantly predicting later mathematics self-concept (see Figure 2). Hence, our results are in accordance with other research in this field (see Marsh & Craven, in press). This relationship held across both samples.

Unlike the relationship between achievement and academic self-concept, we found rather weak evidence for bottom-up or top-down effects relating achievement and global self-esteem in both West and East German samples. Only one path (reflecting a bottom-up effect) was significant. Given the considerable correlations between the three constructs, however, it may be that multicollinearity obscured meaningful effects. To further scrutinize the relationship between self-esteem, academic self-concept, and achievement, we thus ran additional models in which only two of the constructs were included at a time (see Figures 3a–3c).

Figure 2. The dynamics of self-concept: Standardized regression coefficients in the West German (before the slashes) and East German (after the slashes) samples at three measurement points (T1, T2, and T3). T2 residual correlations are not shown. SE = self-esteem, MSC = mathematics self-concept, MACH = mathematics achievement. West/East: $r$(SE-MSC) = -.15***/.13***, $r$(SE-MACH) = .06***/.08***, $r$(SE-MACH) = .04/.03 *$p < .05$. **$p < .01$. ***$p < .001$. 
We first scrutinized the relationship between global self-esteem and mathematics self-concept without controlling for achievement (see Figure 3a). The results of these analyses closely resembled the results already reported, with significantly stronger bottom-up effects in the East German sample than in the West German sample, thus lending further support to the meritocracy principle. Similarly, the reciprocal effects between mathematics self-concept and achievement (Figure 3c) were closely replicated when global self-esteem was excluded from the model.

When only self-esteem and achievement were included in the model (but not mathematics self-concept), four of the eight bottom-up/top-down effects were significantly positive (see Figure 3b), compared with just one significant path coefficient in Model 1a (see Figure 2). Of these four significant path coefficients, three were bottom-up effects. More specifically, achievement consistently predicted later self-esteem in the East German sample, whereas a significant bottom-up effect was only found from T1 to T2 in the West German sample. The one significant but rather moderate top-down effect ($\beta = .05, p < .05$) was found in the East German sample. Hence, as in the Baumeister et al. (2003) review, there was only limited evidence for effects of global self-esteem on achievement.

We next turn to models in which mathematics school grades substituted for the standardized achievement test. The full model incorporating mathematics school grades, mathematics self-concept, and self-esteem showed a reasonably good fit, $\chi^2(546) = 1,244.77$, TLI = .981, RMSEA = .021 (see Model 5a, Table 1). When paths that reflect bottom-up effects were constrained to be equal across the two groups, a significant drop in model fit was observed, $\Delta \chi^2(6) = 38.25, p < .001$. Constraining the top-down effects did not have a significant impact on the model fit, $\Delta \chi^2(4) = 7.01$, ns. The standardized regression coefficients from Model 5a are shown in Table 2. Hence, the moderating effect of the learning environment (East vs. West Germany) on the dynamics between self-esteem and self-concept was unaffected by the substitution of achievement by school grades, providing further support for the meritocracy principle.

Mathematics school grades at T1 significantly predicted self-esteem at T2 (West: $\beta = .07, p < .01$; East: $\beta = .09, p < .001$), but no such effect was found for T2 mathematics school grades. Moreover, no significant effect of self-esteem on later school grades was observed. The impact of self-concept on later school grades was slightly stronger than the impact of self-concept on achievement, but—again—the regression coefficients in West...
Figure 3. Bottom-up and top-down models: Standardized regression coefficients in the West German (before the slashes) and East German (after the slashes) samples. SE = self-esteem, MSC = mathematics self-concept, MACH = mathematics achievement, MGRD = mathematics school grades, GSC = German self-concept, GGRD = German school grades.

*p < .05. **p < .01. ***p < .001.
In the present study, we examined the relationship between achievement, domain-specific academic self-concept, and global self-esteem over time, with a special focus on the moderating role of contextual factors. We first showed that—shortly after reunification—East and West German students experienced their respective learning environments as being markedly different. In line with our hypotheses, East German students were more likely to attribute academic success to effort than West German students, whereas the opposite pattern was found for ingratiation. Hence, East German students articulated a more meritocratic beliefs system. Second, in conformance with the meritocracy principle, the dynamic interrelationship between domain-specific academic self-concepts and self-esteem was markedly moderated by the context. Bottom-up effects (i.e., high academic self-concept leading to higher global self-esteem) were more pronounced in East German students; this relationship held for both mathematics and German grades and self-esteem (Models 7a–7c in Table 1 and Figure 3d). Whereas the reciprocal effects between mathematics school grades and self-esteem (Models 6a–6c in Table 1) and the relationship between mathematics school grades and mathematics self-concept (see Models 6a–6c in Table 1 and Figure 3e) and mathematics school grades and self-esteem, we again found more evidence for bottom-up effects than for top-down effects. Finally, the model in which German grades and German self-concept were included showed both bottom-up and top-down effects.

### Discussion

Despite their shared cultural heritage, the school systems and teaching cultures of East and West Germany before reunification reflected different educational goals. In the East German system, teachers stressed the importance of effort in attaining desirable outcomes, and accurate self-evaluation was an explicit goal (Franz, 1987; Oettingen et al., 1994; Waterkamp, 1987, 1990). This led to an emphasis on social comparisons in East German classrooms, which we expected to prompt bottom-up rather than top-down effects (see Ames, 1992; Wigfield et al., 1996). Because most East

---

**Table 2**

<table>
<thead>
<tr>
<th>Predictors</th>
<th>T2-SE</th>
<th>T2-SC</th>
<th>T2-Grade</th>
<th>T3-SE</th>
<th>T3-SC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mathematics</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-SE</td>
<td>.68***</td>
<td>.09***</td>
<td>.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-SC</td>
<td>.01</td>
<td>.55***</td>
<td>.26***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-Grade</td>
<td>.07**</td>
<td>.05*</td>
<td>.42***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-SE</td>
<td></td>
<td>.62***</td>
<td>.10***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-SC</td>
<td>.00</td>
<td>.48***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-Grade</td>
<td></td>
<td></td>
<td>-.01</td>
<td>.07**</td>
<td></td>
</tr>
<tr>
<td>East</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-SE</td>
<td>.60***</td>
<td>.05</td>
<td>-.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-SC</td>
<td>.14***</td>
<td>.63***</td>
<td>.30***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-Grade</td>
<td>.09***</td>
<td>.08***</td>
<td>.36***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-SE</td>
<td></td>
<td>.57***</td>
<td>.07**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-SC</td>
<td>.17***</td>
<td>.57***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-Grade</td>
<td></td>
<td></td>
<td>-.04</td>
<td>.09***</td>
<td></td>
</tr>
</tbody>
</table>

**German**

<table>
<thead>
<tr>
<th>Predictors</th>
<th>T2-SE</th>
<th>T2-SC</th>
<th>T2-Grade</th>
<th>T3-SE</th>
<th>T3-SC</th>
</tr>
</thead>
<tbody>
<tr>
<td>West</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-SE</td>
<td>.69***</td>
<td>.18***</td>
<td>.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-SC</td>
<td>-.04</td>
<td>.38***</td>
<td>.14***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-Grade</td>
<td>.09***</td>
<td>.17***</td>
<td>.50***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-SE</td>
<td></td>
<td>.59***</td>
<td>.08**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-SC</td>
<td>.06*</td>
<td>.45***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-Grade</td>
<td>.07*</td>
<td>.11***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>East</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-SE</td>
<td>.62***</td>
<td>.11***</td>
<td>-.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-SC</td>
<td>.10**</td>
<td>.55***</td>
<td>.18***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1-Grade</td>
<td>.02</td>
<td>.01</td>
<td>.43***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-SE</td>
<td></td>
<td>.58***</td>
<td>.07*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-SC</td>
<td>.14***</td>
<td>.61***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2-Grade</td>
<td></td>
<td>.04</td>
<td>.09***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. T1, T2, T3 = measurement points; SE = global self-esteem; SC = mathematics/German self-concept; Grade = mathematics/German grade. ***p < .001; **p < .01; *p < .05.

($\beta = .26, p < .001$) and East ($\beta = .30, p < .001$) were comparable. Additional models were specified to examine the dynamic relationship between mathematics school grades and mathematics self-concept (see Models 6a–6c in Table 1 and Figure 3e) and mathematics school grades and self-esteem (Models 7a–7c in Table 1 and Figure 3d). Whereas the reciprocal effects between mathematics self-concept and mathematics school grades were unaffected by the omission of self-esteem, stronger bottom-up effects between T1 mathematics grades and T2 self-esteem emerged once mathematics self-concept was dropped from the model (West: $\beta = .08, p < .001$; East: $\beta = .14, p < .001$).

Finally, we performed a similar set of analyses with German school grades and German self-concept in place of mathematics grades and mathematics self-concept. The most comprehensive model including all three constructs showed a reasonably good fit, $\chi^2(546) = 1,260.67$, TLI = .976, RMSEA = .022 (see Model 8a, Table 1). When paths that reflect bottom-up effects were constrained to be equal across the two groups, a significant drop in model fit was again observed, $\Delta \chi^2(6) = 42.93, p < .001$. Constraining the top-down effects did not significantly worsen the model fit, $\Delta \chi^2(4) = 7.28, ns$. The standardized regression coefficients from Model 8a are reported in the lower part of Table 2.
German teachers were allowed to stay in their jobs after reunification, it is reasonable to assume that this kind of teaching culture continued to prevail to a certain degree, even after the transformation of the East German system.

Our empirical data indeed show some marked differences between the East and West German learning environments. East German students attributed academic success more strongly to effort and less strongly to ingratiation than their West German counterparts. These differences, which amounted to more than half a standard deviation for both variables, are in accordance with our theoretically derived hypotheses and are based on a large, representative sample of students. Hence, these data provide particularly strong evidence for differences in the respective learning cultures. Moreover, the context proved to affect the relationship between the two indicators of the learning environment. Whereas a positive correlation between effort and ingratiation was found in West German students, the correlation was negative for East German students. Hence, East German students seem to have acquired a more differentiated beliefs system regarding the causes of academic success.

Research in the field of developmental and educational psychology indicates that differences in the learning environment have an impact on various psychological outcomes (see Pintrich, 2003; Wigfield et al., 1996). The more supportive and interested in individual progress a teacher is perceived to be, for example, the higher the academic self-concept of his or her students (Lüdtke et al., 2005; Wigfield et al., 1996). At the same time, however, very few studies have been conducted on whether causal dynamics are moderated by different learning contexts. Hence, in exploring the meritocracy principle, the present study extended prior research on the effects of different learning environments. Moreover, our study adds to the literature on the relative benefits of classrooms in which student learning (task goal structure) or student performance (ego goal structure) are emphasized (e.g., Ames, 1992; Pintrich, 2003). Meritocratic environments entail aspects of both task orientation (emphasis on effort) and ego orientation (emphasis on social comparison). Hence, meritocratic environments are, to a certain extent, learning environments that are high on both task and ego orientation.

The Meritocracy Principle

The meritocracy principle holds that a student’s evaluation of his or her achievement in academic domains (academic self-concept) has a more pronounced impact on his or her global self-evaluation (self-esteem) in meritocratic learning environments. Accordingly, we expected to find stronger bottom-up effects in East than in West German students.

Our empirical results confirm the assumptions of the meritocracy principle. In the East German sample, the paths from domain-specific self-concepts to global self-esteem were consistently and significantly positive, independent of the domain (mathematics or German) and at both time intervals (beginning to middle of Grade 7; middle to end of Grade 7) under consideration. The inclusion of different achievement indicators (achievement tests; school grades) only had a tangential impact on the size of these effects. At a casual glance, the bottom-up effects may seem to be of moderate size ($11 < \beta < .17$). However, beta coefficients for cross-lagged effects in the range of .10–.15 are common in real-world, nonexperimental, longitudinal research and can be considered meaningful (e.g., Marsh & Craven, in press; Roberts, Caspi, & Moffitt, 2003). When interpreting these cross-lagged effects, one should bear in mind that our analyses controlled for the stability of self-esteem and the effects of achievement indicators. Moreover, change in self-esteem is, of course, multiply determined (Ahadi & Diener, 1989). Academic achievement and academic self-concepts are just one force in the dynamics of self-concept; self-esteem is also subject to influences from the family and peer contexts, for example (e.g., Harter, 1998). It should also be noted that cross-lagged effects are potentially cumulative over time; given a beta coefficient in the range of .10–.15 for a 6-month period, the specific effect may be quite substantial if it continued over longer periods of time (Neyer & Asendorf, 2001; Prentice & Miller, 1992).

In contrast to the pattern of results for the East German sample, support for bottom-up effects from academic self-concept to self-esteem was very low in the West German sample. Only one bottom-up effect reached the level of significance (T2 German self-concept on T3 self-esteem), and this effect was small in size ($\beta = .06, p < .05$). Top-down effects (from self-esteem to academic self-concept) were predominant in the West German sample.

One central finding of the present study is thus the empirical support it provides for the learning environment as moderator of the development of self-concept. Although the importance of the learning environment has been stressed by many researchers (e.g., Ames, 1992; Wigfield et al., 1996), to our knowledge, this is the first nonexperimental study to document significant effects of the learning environment on the dynamics between global self-esteem and domain-specific academic self-concept.

With its large-scale, longitudinal, multivariate design and representative sample of students, our study presents a particularly strong basis for testing the predictions of the meritocracy principle. Of course, these predictions are not restricted to the German school system. We expect that additional support for the meritocracy principle will be found in future cross-cultural studies within or between different countries. Researchers seeking to examine the meritocracy principle should design studies that fulfill at least three prerequisites: (a) have a longitudinal design, (b) include measures that can be used to describe the level of perceived meritocracy, and (c) administer identical measures of self-esteem and academic self-concept at all measurement points. Inclusion of objective achievement indicators would clearly strengthen such cross-cultural studies even further.4

Does High Self-Esteem Lead to Higher Academic Achievement?

The recent review by Baumeister et al. (2003) on the effects of self-esteem on achievement has attracted considerable attention from the scientific and nonscientific public because its findings

---

4 Unfortunately, the comparative international studies on educational attainment that have become increasingly prevalent in recent years (e.g., Organisation for Economic Cooperation and Development, 2001) do not allow the meritocracy principle to be tested because the studies are typically cross-sectional in design.
seem to run counter to the popular belief that a positive self-view leads to higher achievement. However, the review was based on very few studies—there are only a handful of data sets that allow in-depth study of this relationship over time (see Marsh et al., 1999, for a related discussion of studies on achievement and academic self-concept). The present study, with its large, representative sample, consideration of school grades and standardized achievement tests as alternative achievement indicators, and repeated administration of self-esteem items, makes a substantial contribution to this narrow literature.

In general, our results are in line with the conclusions of Baumeister et al. (2003), indicating that high self-esteem is not a strong predictor of later achievement. When we restricted our analyses to the reciprocal relationship between self-esteem and achievement, we found some significant top-down effects, but these effects were generally rather small (.05 < β < .07). When we controlled for domain-specific academic self-concept, we no longer found any direct effects of self-esteem. Hence, the effects of self-esteem on achievement—if they exist at all—appear to be mediated by domain-specific academic self-concepts. Although our results generally correspond with the findings of Baumeister et al. (2003), we would like to stress that their review was restricted to the relationship between global self-esteem and achievement. Given the great interest in global self-esteem that can be observed in social and personality psychology as well as in nonscientific contexts, this restriction is understandable. However, it is important to highlight that the small effects observed of self-esteem on achievement do not imply that self-views have no impact on achievement. In recent years, most research in the educational and developmental field has concentrated on the relationship between achievement and domain-specific academic self-concepts (rather than self-esteem), and these studies generally support the view that a high academic self-concept leads to higher achievement (see Marsh & Craven, in press; Valentine et al., 2004). Our study lends further support to this assumption. We found considerable effects of academic self-concept on achievement. Again, this held for both standardized achievement indicators and school grades and was replicated across mathematics and German and across East and West German students, and the regression coefficients were of considerable magnitude (.07 < β < .30). Taken together, the data indicate that domain-specific academic self-concept is a significant predictor of later achievement.

Limitations and Future Research

The present study makes a substantial contribution to the debate on the relationship between achievement, domain-specific academic self-concept, and self-esteem. More important, it is the first study to show that the relationship between the global self-view and the evaluation of accomplishments in the academic domain is considerably moderated by the learning context. With its large, representative sample of 7th graders examined three times over the course of a school year and its set of high-quality achievement and self-concept indicators, the present study was particularly well suited to addressing our research questions.

At the same time, however, we would like to point to some limitations of the present study. First and foremost, our study has a limitation common to practically all real-world, nonexperimental research: the possibility of third-variable explanations. For instance, we argued that bottom-up effects would be more dominant in meritocratic learning environments and were able to confirm the postulated moderating effects on the dynamics of self-concept. An outside observer might agree with our contention that the observed moderator effects are of high theoretical and practical interest but disagree with our explanation of this effect. The East and West German learning environments might well differ on a number of characteristics apart from their meritocratic orientation, and these other characteristics might in fact have been responsible for the moderator effects that we found for the dynamics among self-esteem and domain-specific self-concept. Unfortunately, there is no ideal solution to this third-variable problem in this study (or indeed in nonexperimental, real-world studies in general). However, because the meritocracy principle was theoretically derived and is postulated to hold across different cultures, the main postulates of our study can be easily scrutinized in cross-cultural research, providing that appropriate data sets are available. We hope that future studies will investigate the meritocracy principle in a variety of contexts, using experimental and nonexperimental designs.

As a further limitation, we note that our study was restricted to 7th graders. Although we did not hypothesize that the relationship between achievement, academic self-concept, and self-esteem is moderated by age, we cannot rule out a priori such a possibility. Moreover, the time frame of our study was just 1 year. The rather high overall stability of self-esteem found in our study as well as in other studies (e.g., Trzesniewski, Donnellan, & Robins, 2003) suggests that effects of the environment on change in self-esteem may take some time. It is also quite possible that top-down and bottom-up processes play out differently over short versus medium versus long time periods. On the basis of the findings of experimental studies by Brown (1993), who reported top-down effects of considerable size, one might speculate that top-down effects are more pronounced over short time periods but diminish over periods of several months, as in the present study. Similarly, as suggested by Shavelson et al. (1976), repeated feedback on achievement might be necessary to change more global self-evaluations; hence, it may take longer than 1 school year for bottom-up effects to develop fully. Studies covering a time span of more than 1 school year are needed here.

Our study was also restricted in the sense that we concentrated on academic achievement and academic self-concepts and did not include any nonacademic domains (e.g., social status and social self-concept). Cross-cultural studies that included both academic and nonacademic self-concepts and examined their relative predictive power for global self-esteem would make a highly valuable contribution to the literature.

Finally, from a practical point of view, our study does not allow conclusions to be drawn on whether meritocratic or ego-protective learning environments are preferable. Although we found stronger bottom-up effects in the meritocratic East German learning environment, this does not mean that this learning context is preferable over time with respect to the development of self-esteem and self-concept. To judge the relative benefits of meritocratic versus ego-protective environments, in addition to their effects on self-concept and self-esteem, researchers will have to examine the long-term effects of these environments on a number of additional outcome variables, such as achievement, interest, and global school adjustment.


Received May 27, 2003

Revision received September 30, 2005

Accepted October 7, 2005