UNIVERSITY of CALIFORNIA PRESS | Collabra: Psychology

Schneider, S., Lösch, T., Schneider, D., & Schütz, A. (2022). How Self-Concept, Competence, and Their Fit or Misfit Predict Educational Achievement, Well-Being, and Social Relationships in the School Context. *Collabra: Psychology*, *8*(1). https://doi.org/10.1525/collabra.37154

Personality Psychology

How Self-Concept, Competence, and Their Fit or Misfit Predict Educational Achievement, Well-Being, and Social Relationships in the School Context

Sarah Schneider[1], Thomas Lösch[2], Daniel Schneider[1], Astrid Schütz[1] [a]

[1] Psychology, University of Bamberg, Bamberg, Germany, [2] DIPF | Leibniz-Institute for Research and Information in Education, Frankfurt, Germany

Keywords: self-concept, competence, achievement, well-being, social relationships, multilevel, panel data, NEPS

https://doi.org/10.1525/collabra.37154

Collabra: Psychology

Vol. 8, Issue 1, 2022

During adolescence, what is more important for educational achievement, well-being, and the formation of positive social relationships: being competent, having positive thoughts about oneself, or a complex relationship between the two? There has been a long-standing debate in psychology on the effects of accurate and biased self-perceptions, and sophisticated ways of modeling the effects of self-perception, competence, and their interplay have recently been suggested. But recent research has focused on adults and has not taken reference effects into account. The present preregistered study used a large German sample of students (*N* = 6,086 students in 559 classes) in Grade 5 (mean age = 10.55 years, *SD* = 0.64) with data from the National Educational Panel Study (NEPS). We tested the effects of academic self-concept and competence in math and reading on outcomes pertaining to achievement, well-being, and social relationships up to 4 years later and identified the best fitting hypotheses through a model fit comparison. In contrast to previous studies, we took the frame of reference for students' self-concept into account by controlling for class-level effects of self-concept and competence in a multilevel analysis. Results showed that educational achievement was best explained by the complex interplay of self-concept and competence, where competence was the stronger predictor. By contrast, self-concept was a stronger predictor of well-being than competence was. For social relationships, results were less clear and differed by the specific outcome variables that were used. Overall, in the school context, self-concept and competence per se seem to be more predictive of future outcomes than their fit or misfit.

People's perceptions of themselves and their own competence (e.g., self-concept) and the fit or misfit of those self-perceptions with external criteria (e.g., actual competence) have been found to predict indicators of achievement (Chung et al., 2016; Paschke et al., 2020), well-being (Bonanno et al., 2005; Colvin et al., 1995; Paulhus, 1998; Schimmack & Kim, 2020), and social relationships (Anderson et al., 2012; Brown, 1986; Dufner et al., 2019; Robins & Beer, 2001). The current study focuses on self-perceptions, specifically people's academic self-concepts. Academic self-concepts refer to a person's domain-specific self-perceptions of competence (Arens et al., 2011). One can think of oneself in a realistic or biased way (John & Robins, 1994). On the one hand, one can have an accurate perception of one's actual competence (e.g., showing fit with an external criterion), but on the other hand, one can overestimate or underestimate oneself (e.g., showing misfit with an external criterion).

Researchers in social, personality, and educational psychology have investigated how the fit or misfit between self-concept and actual competence predicts outcomes, such as educational achievement, well-being, and social relationships (e.g., Baumeister et al., 2003; Dunlosky & Lipko, 2007), but findings have been mixed (Colvin & Block, 1994; Goorin & Bonanno, 2009; Kwan et al., 2004; Robins & Beer, 2001; Taylor & Brown, 1988). However, some studies suggested that these contradictory findings were the result of different methodological choices that often made it difficult to separate effects of self-concept from effects of fit (Kwan et al., 2004). Recent strategies have helped to disentangle and clarify such effects (Edwards, 2002; Humberg et al., 2018) by modeling the effects of self-concept, competence, and their relationship as multiple components in a polynomial regression model. By also specifying several different regression models and testing the models against each other, researchers have been able to identify which rela-

a Corresponding author: Astrid Schütz, University of Bamberg, Markusplatz 3, 96047 Bamberg, Germany
Email: astrid.schuetz@uni-bamberg.de

tionships between self-concept and actual competence best explain different life outcomes in adults (Humberg et al., 2019).

The modeling of multiple components of self-concept and external criteria as well as their fit or misfit has only rarely been applied to study effects on educational achievement (Paschke et al., 2020), well-being, or social relationships (Humberg et al., 2018, 2019; Schimmack & Kim, 2020). Nevertheless, all three outcomes are relevant life outcomes as originally formulated by Taylor and Brown (1988). Furthermore, all three outcomes are especially relevant in the school context where children face a number of developmental challenges. Thus, a systematic model comparison that includes a broad array of outcomes in the school context is warranted. Furthermore, contextual factors should be taken into account because self-ratings are typically made in relation to a comparison group (Marsh, 1987).

The present study tests how self-concept, competence, and their relationship affect achievement, well-being, and social relationships in a large national sample of school students in Grade 5 using a model comparison approach. Furthermore, the use of multilevel analyses ensures that reference effects at the class level are controlled for (Snijders & Bosker, 2011).

Educational Achievement, Well-Being, and Social Relationships in the School Context

The transition from childhood to adulthood during adolescence is challenging. Adolescents often experience a lot of pressure and expectations regarding their educational achievement (Pinquart & Ebeling, 2020) while they are exploring their own identities (Harter, 2006) and trying to form and maintain healthy social relationships (Pfeifer & Berkman, 2018; Schwartz et al., 2006). Well-being and social relationships are often unstable during adolescence (Harris & Orth, 2020; Harter, 2006; Wagner et al., 2018). Regarding social relationships, in comparison with adulthood, adolescence is a life stage in which individuals have a multitude of social contacts (Lansford et al., 1998; Luong et al., 2011) and where social networks with peers become increasingly important (Giordano, 2003). Moreover, educational achievement, well-being, and social relationships in adolescence have been found to predict later life success, mental health, and relationships (Harris & Orth, 2020; Roberts et al., 2007; Steinmayr et al., 2014). For example, educational achievement during adolescence is positively linked to early career adaptability (Negru-Subtirica & Pop, 2016) and chances for career promotion in adult life (Ng et al., 2005; Steinmayr et al., 2014). Finally, bonding and forming social relationships with peers can aid the development of interpersonal skills (Hansen et al., 1992) and enhance mental health in midlife (Hightower, 1990).

Educational achievement refers to general or domain-specific knowledge that students obtain in their school life and everyday life (e.g., Messick, 1984). Commonly used indicators of educational achievement are school grades and standardized *competence tests* that were designed to measure general knowledge that students obtain in their school life and everyday life (e.g., Heckman & Kautz, 2012). Competence tests are used to assess educational achievement more objectively than grades, which also reflect other characteristics (e.g., what teachers value in students' school work; Brookhart et al., 2016; Casillas et al., 2012).

Well-being refers to how people experience quality of life on emotional and cognitive levels (Diener, 1984). It includes a low level of *stress*, as stress is one of the most important indicators of both mental and physical health (Backé et al., 2012; Siddique & D'Arcy, 1984). *Self-esteem* is a global evaluation of one' s own worth (Baumeister, 1999; Rosenberg, 2016). Quality of life at emotional and cognitive levels is often measured as *life satisfaction* (Diener, 1984; Diener et al., 2018), which describes how people experience various areas of their lives, such as work life, family, and friends (Çivitci & Çivitci, 2009; Robert A. Cummins, 2005).

Social relationships in the school context are important for adolescents: Students who are socially well-integrated and engage in positive relationships with their peers are more likely to take advantage of social and learning opportunities (Birch & Ladd, 1996) and show better physical health as adults (Allen et al., 2015). Social relationships have been found to be predicted by self-concept and competence. Whereas children with more positive self-concepts have tended to be more popular (Jackson & Bracken, 1998; Verschueren et al., 2001), there have been mixed results on whether higher achieving children are more (Kiuru et al., 2015; Vannatta et al., 2009) or less popular (e.g., Bellmore, 2011). Social relationships during adolescence are evidenced by various aspects, such as the quality and *quantity of friendships* (Bukowski & Hoza, 1989) and *social integration* in a class as shown by group acceptance (Birch & Ladd, 1996) versus *peer problems* or harassment within that group (Goodman, 1997).

Conceptualizations of Self-Concept and Its Fit or Misfit With External Criteria

Self-concept, competence, and their fit or misfit with each other as predictors of different life outcomes have been investigated in several fields of psychological research, for example, in social and personality psychology (e.g., Baumeister et al., 2003) or educational psychology (Chiu & Klassen, 2009, 2010).

Self-concept describes one's view of oneself or how the mind perceives itself (e.g., James, 1890; Schütz & Baumeister, 2017). Global self-concept is understood as a multidimensional structure that includes nonacademic and academic components, whereas academic self-concepts refer to domain-specific self-perceptions in various competence domains, such as mathematics or reading (Arens et al., 2011; Marsh & Shavelson, 1985). Researchers have argued that academic self-concept is already relevant in and of itself (Marsh & Craven, 2006; Marsh & Martin, 2011), but additionally, self-concept has been shown to predict a range of future outcomes, such as academic engagement and achievement (e.g., Guo et al., 2016) or educational choices (e.g., Guo et al., 2015).

It has been argued that self-concept's fit or misfit with an external criterion is a relevant predictor of multiple outcomes (e.g., Dufner et al., 2019). Misfit is understood as self-perception that deviates from a reference point in an external criterion (e.g., actual competence) and is described as "illusory self-perception" in the case of positive misfit (Baumeister, 1989; Dufner et al., 2019). Such misfit can occur in both directions, such that it can be more positive (self-enhancement) or more negative (self-effacement) than the criterion. Self-perception fit is understood as self-perception that does not deviate from an external criterion.

There are several explanations for why there may be misfit between self-perception and external criteria. For example, in social and personality psychology, positive illusions about the self are often understood as guided by a motive to increase well-being or inflate one's self-worth (Leary, 2007; Schütz & Baumeister, 2017). This motive entails optimism in terms of one's control over events and the future as well as self-serving attributions in which success is attributed to the self and failure is attributed to chance or circumstances (Mezulis et al., 2004; Sedikides & Gregg, 2008). By contrast, negative misfit or self-effacement describes a self-critical motive by which one elaborates on and emphasizes negative aspects of the self (Kitayama & Karasawa, 1997). It is often aimed at maintaining or protecting social relationships instead of distinguishing oneself from others (Heine et al., 1999; Heine & Lehman, 1995; O'Mara et al., 2012). In educational psychology, it is often argued that misfit is the result of a lack of the (meta-)cognitive ability or skill to accurately perceive one's own competence (e.g., Dunlosky & Lipko, 2007).

Modeling the Effects of Self-Concept, Competence, and Their Fit or Misfit

Different operationalizations of the fit or misfit between self-concept and external criteria have led to inconsistent results regarding their effects (e.g., Kwan et al., 2004). Many studies have relied on difference or residual scores (for a meta-analytical review, see Dufner et al., 2019), but there are limits to such an approach. Difference and residual score approaches use a single score to represent fit and misfit. A difference score is calculated as the difference between self-concept and the external criterion. A residual score is calculated as the residuum of self-concept regressed on the external criterion (for an overview, see Barranti et al., 2017; Edwards & Parry, 1993). However, the use of a difference or residual score as a single indicator of the misfit between self-concept and an external criterion results in a loss of information because the levels of self-concept and the external criterion cannot be preserved (Asendorpf & Ostendorf, 1998; Barranti et al., 2017; Humberg et al., 2019). For example, having a score of 1 for a biased self-concept does not provide information about the levels of self-concept and the external criterion at which this discrepancy occurred. However, the levels at which misfit occurs are likely to have an impact on outcomes (Humberg et al., 2019).

Polynomial regression and response surface methodology offer useful solutions to this problem (Edwards, 1994b, 2002). In polynomial regression, the outcome is regressed on the linear, curvilinear, and product terms of several predictors (e.g., Edwards, 1994a; Edwards & Parry, 1993). Response surface methodology (also called response surface analysis; RSA) provides a precise description and meaningful interpretation of a three-dimensional surface that corresponds to the respective polynomial regression model (Box & Draper, 1987; Edwards & Parry, 1993; Khuri & Cornell, 1987). RSA models the first- and second-order terms in a pair of predictors and their association with the outcome in a three-dimensional response surface (Barranti et al., 2017). Therefore, the effect of fit and misfit can be evaluated in relation to the levels of the predictors.

Evidence of the Effects of Fit and Misfit between Self-Perception and Competence

A recent review summarized comprehensive evidence of the effects of the fit or misfit of self-perception with external criteria (Dufner et al., 2019). The review addressed the inconsistencies that have occurred in the results of previous research and pointed out that the inconsistencies were due to the use of insufficient methodologies (e.g., relying on difference scores as measures of misfit). The review found that for well-being, the positive misfit between self-perception and an external criterion was beneficial overall. For social relationships, the effect of the misfit between self-perception and an external criterion seemed to depend on acquaintance (positive misfit had a positive effect at zero acquaintance that diminished with longer acquaintance) and the domain of the misfit, where agentic and communal misfit had positive effects on their own.

Unfortunately, research that has relied on difference or residual scores is not able to adequately inform the present questions. A more suitable approach is polynomial regression and subsequent model comparisons, which use not only information such as the level of fit or misfit but also the respective levels of both self-concept and competence—an aspect that is not considered when difference or residual scores are used. Thus, in the following, we refer to studies that used a polynomial regression-based approach to study the effects of self-perception and competence. Studies that did not rely on polynomial regression may have provided diverging results because they used a different methodology.

Humberg et al. (2019) tested the effects of self-perception, intellectual ability, and their fit or misfit on outcomes of psychological adjustment in a sample of young adults. In summary, their study found especially beneficial effects of self-perception on well-being, whereas for social relationships, actual intellectual ability also had a beneficial effect on peer-rated agentic outcomes. However, the misfit between self-perception and actual intellectual ability explained the data best only when the outcomes of agency and communion were self-rated.

Another study that used multiple components for both self-concept and external criteria and scores for their relationship tested effects on academic achievement in a sample of school students in Grade 10 (Paschke et al., 2020). The results provided evidence of beneficial effects of positive self-perception and competence on academic achievement, but it found no support for the misfit (Paschke et al., 2020).

However, previous research on the effects of self-perception, competence, and the extent to which self-perception is accurate (i.e., shows fit with competence) has not taken into account frame-of-reference effects. A frame of reference is defined as the context in which an individual's perception occurs (Bullock & Trombley, 1999). In the school context, the most prominent reference effect is the Big-Fish-Little-Pond Effect (BFLPE; Marsh, 1987). According to the BFLPE, a student in a class with a low class-average reading competence will have a more positive reading self-concept than the same student in a class with a high class-average reading competence. The frame of reference—in this case, the competence level of the whole class—affects the way a student perceives them self. If the reference effect is not taken into account, one could falsely assume that the student with a positive self-perception, who has, for example, a mediocre actual level of competence, self-enhances. However, the student's (overly) positive self-perception might occur only in a class with a mediocre class-average competence level and not in a class with a high class-average competence level and may therefore be due to the frame of reference. In order to differentiate between reference effects and effects of misfit between self-concept and competence, both should be considered in the modeling strategy, such as in a multilevel model.

Model Fit Comparisons to Identify the Best-Fitting Hypotheses

A number of hypotheses exist on the effects of self-concept, competence, and their fit or misfit with each other. To identify which of these hypotheses is most likely to predict outcomes in school, they can be compared on the basis of their model fit (e.g., Newland, 2019). To specify the hypotheses correctly, polynomial regression techniques can be adapted.

Polynomial regression allows for the specification of a more comprehensive and precise set of hypotheses that describe the effects of self-concept, competence, as well as their fit and misfit. Humberg et al. (2019) summarized and specified different hypotheses on the effects of self-perception, competence, and their relationship. We adapted the hypotheses and respective models to our research questions (see Table 1) and tested them with our data.

The set of hypotheses begins with hypotheses that describe positive linear effects of only self-concept or competence as well as both self-concept and competence (*Beneficial Positive Self-Concept Only Hypothesis, Beneficial Competence Only Hypothesis, Beneficial PSC & Competence Hypothesis*). The next two describe negative linear effects of self-concept and competence (*Detrimental PSC Only Hypothesis*, *Detrimental PSC & Competence Hypothesis*), followed by two hypotheses that focus on the misfit of self-concept and competence in the form of a positive effect of positive misfit (*Beneficial Positive Misfit & Detrimental Negative Misfit Hypothesis*) or a negative effect of positive misfit (*Detrimental Positive Misfit & Beneficial Negative Misfit Hypothesis*). The *Accurate Self-Concept Hypothesis* posits a positive effect of a fit between self-concept and competence, whereas the *Optimal Margin Hypothesis* assumes a positive effect of positive misfit up to an optimal level. Two hypotheses describe a positive curvilinear effect that decreases at a certain saddle point for each measurement of self-concept and competence (*Curvilinear PSC Hypothesis, Curvilinear Competence Hypothesis*). The *Interaction Hypothesis* proposes a positive interaction effect of self-concept and competence, meaning that the effect of either predictor will be more positive when the other predictor is at a higher level. Finally, the *Global Hypothesis* includes linear and curvilinear terms of each predictor as well as their product. All hypotheses are tested against the *Null Hypothesis*, which states that self-perception and competence have no relationship with the outcomes at all.

The approach of modeling multiple components with polynomial regression has been successfully used to investigate other fit hypotheses (Deventer et al., 2019; Förster et al., 2022; Mielke et al., 2021; Mota et al., 2019). We used an information-theoretic approach with the second-order Akaike Information Criterion (AICc) as an indicator of model fit (Sugiura, 1978) and aimed to select the model that explained the data best according to a model fit comparison. Information-theoretic model selection is an analytic approach that builds on Maximum Likelihood estimates, such as the Akaike or Bayes Information Criterion. By using a measure of fit to compare a set of model candidates, one can select the model that has the highest probability of explaining the data best (Newland, 2019; Rost, 2004).

On the basis of previous results on adults, we expected to find distinct results for the three outcomes. For outcomes regarding educational achievement, we expected beneficial effects of self-concept (i.e., our measure of self-perception) as well as competence. For outcomes regarding well-being, we expected that, similar to previous research (Dufner et al., 2019; Humberg et al., 2019), positive self-concept would be the best predictor. For outcomes regarding social relationships, we expected the data to be explained best by the models including beneficial effects of self-concept and competence and positive misfit.

The Present Investigation

In the present investigation, we sought to analyze how self-concept, competence, and their fit or misfit with each other would predict future achievement, well-being, and social relationships. To this end, we used polynomial regression models and RSA (Edwards, 2002; Edwards & Parry, 1993), and we took the Table 1 hypotheses that Humberg et al. (2019) had tested in a sample of adults, and we tested the same hypotheses in a sample of adolescents. Following an information-theoretic approach, we aimed to determine which of the hypotheses would fit the data best according to a model fit comparison. As outcome categories we used educational achievement, well-being, and social relationships, representing the major life domains as originally discussed by Taylor and Brown (1988). We looked at three academic domains (math, reading, and general academic skills) and used self-concept and competence indicators as predictors. Finally, we took into account reference effects at the class level by using multilevel modeling (Snijders & Bosker, 2011). We used a large sample of school students and analyzed outcomes over the course of several years with data from the National Educational Panel Study (NEPS; Blossfeld & von Maurice, 2011) in Germany. The design of our study supported its external validity, as it was embedded in the applied context of the educational domain.

Method

Preregistration

We preregistered our analytic procedure prior to obtaining and analyzing the data. Our preregistration can be found at https://osf.io/cn67w.[1] We followed our preregistered plan the best we could for our analyses. At some points, we had to deviate from our preregistered plan due to nonconvergence in the models, meaning that we had initially specified too many parameters in the estimation by allowing all predictors across all domains (math, reading, general) to be correlated in each model, which resulted in models that could not be estimated. We then used models with fewer parameters, see Footnote 2.

Sample

We used data from the cohort of school students from the National Educational Panel Study (NEPS; Blossfeld & von Maurice, 2011). The NEPS is a large-scale German panel study conducted by the Leibniz Institute for Educational Trajectories consisting of multiple cohorts with the major goal of tracking individuals' development across the entire life span (Blossfeld & von Maurice, 2011). The data we analyzed came from the scientific use file (SUF) version 8.0.0 (Weinert & NEPS, National Educational Panel Study, 2019).

For the present study, we used a sample that was in Grade 5 at the first wave of measurement in the 2010/2011 school year. In most federal states in Germany, Grade 5 is the first year of tracked secondary education. Students from all tracks are included in the NEPS. We aimed to gather a representative sample of the whole student population in Grade 5 in Germany. Compared with the national school population, lower secondary schools (Hauptschulen) were slightly underrepresented, whereas high schools were slightly overrepresented (IEA Data Processing and Research Center, 2012). Furthermore, in Wave 1, students in special needs schools had a higher propensity to participate, whereas students with a native language other than German had a lower propensity to participate (see Steinhauer & Zinn, 2016).

Measurements took place once per school year (Waves 1 to 3 and 5 were measured from November to January, Wave 4 was measured from November to May), except for Wave 6, which was measured in the same school year (2014/2015) as Wave 5 but in different months (April 2015 to July 2015). Students worked on the competence tests and questionnaires in their regular school classes, or their parents were interviewed via telephone (computer assisted telephone interview). We used data from Wave 1 for the predictor variables and data from Waves 3 (school year 2012/2013) to 5 (school year 2014/2015) and 6 (school year 2014/2015) for the outcome variables (see Table 2 for measurement periods).

Overall, our sample included 6,112 students in 564 classes. Of these, 2,950 students (48.6%) were female, 3,125 students (51.4%) were male, and there was no information on gender for 37 students. In Wave 1, students had a mean age of 10.55 years (*SD* = 0.64, there was no information on age for 11 students). In Wave 1, students reported their birth country. Of the 6,112 students in the current sample, 5,391 (88.2%) said their birth country was Germany, 117 (1.9%) said they were from another European country, 81 (1.3%) said they were born in a state of the former Soviet Union, and 48 (0.8%) said they were born in another country that was not part of Europe. For 475 (7.8%), no such information was available. For further details, see Olczyk et al. (2016). The sample included students from the different school tracks. Of the 6,112 students, 336 (5.5%) were in primary school; 745 (12.2%) were in lower secondary education ("Hauptschule," which qualifies as elementary education); 526 (6.6%) were in schools with several courses of education; 1,175 (19.2%) were in "Realschule," which qualifies as intermediate secondary education; 328 (5.4%) were in integrated comprehensive schools; 2,415 (39.5%) were in high school, qualifying for the general qualification for university entrance; and 587 (9.6%) were in special needs schools. After excluding students who had missing values on all analyzed variables, the sample consisted of 6,086 students in 559 classes. However, in some specific models, students had missing data on all the variables we used. Therefore, we could use between 6,081 and 6,086 cases depending on the specific model.

Measures

To study the effects of self-concept, competence, and their relationship on educational achievement, well-being, and social relationships, we used a number of measures as predictors and outcomes. As predictors, we used academic self-concept in math and reading as well as general academic self-concept as measures of self-perception and performance on competence tests as a measure of competence. As outcomes, we used (a) performance on competence tests in math and reading as measures of educational achievement, (b) self-esteem, stress, and life satisfaction as measures of well-being, and (c) peer problems (reported by students and their parents) as well as parent reports of number of friends and social integration as measures of social relationships.

In the NEPS, not all measures are administered in each wave. Thus, all outcomes were not measured in the same wave. We used Waves 3 to 6 for the outcomes and Wave 1

1 Please note that the time-stamped preregistration was unfortunately submitted with the authors' names visible. The manuscript is thereby linked to an anonymized version of the preregistration. We gave the handling editor the link to the identifiable and timestamped preregistration.

Table 1. Hypotheses and Statistical Models

Hypothesis	Verbal description	Statistical model
1. Beneficial Positive Self-Concept (PSC) Only Hypothesis	For two students, the outcome Z is higher for the person with the higher self-concept S.	Beneficial Positive Self-Concept (PSC) Only Model : $Z = b_0 + b_1S + b_2C + \varepsilon$ with $b_1 > 0, b_2 = 0$
2. Beneficial Competence Only Hypothesis	For two students, the outcome Z is higher for the person with the higher competence C.	Beneficial Competence Only Model: $Z = b_0 + b_1S + b_2C + \varepsilon$ with $b_1 = 0, b_2 > 0$
3. Beneficial PSC & Competence Hypothesis	For two students with the same value (level) of competence C, the outcome Z is higher for the person with the higher self-concept S (and vice versa).	Beneficial PSC & Competence Model: $Z = b_0 + b_1S + b_2C + \varepsilon$ with $b_1 > 0, b_2 > 0$
4. Detrimental PSC Only Hypothesis	For two students, the outcome Z is lower for the person with the higher self-concept S.	Detrimental PSC Only Model: $Z = b_0 + b_1S + b_2C + \varepsilon$ with $b_1 < 0, b_2 = 0$
5. Detrimental PSC & Competence Hypothesis	For two students with the same value (level) of competence C, the outcome Z is lower for the person with the higher self-concept S (and vice versa).	Detrimental PSC & Competence Model: $Z = b_0 + b_1S + b_2C + \varepsilon$ with $b_1 < 0, b_2 < 0$
6. Beneficial Positive Misfit & Detrimental Negative Misfit Hypothesis	The higher the positive discrepancy between self-concept S and competence C of a student, the higher is the outcome Z.	Beneficial Positive Misfit & Detrimental Negative Misfit Model: $Z = b_0 + b_1S + b_2C + \varepsilon$ with $b_1 > 0, b_2 < 0$
7. Detrimental Positive Misfit & Beneficial Negative Misfit Hypothesis	The higher the positive discrepancy between self-concept S and competence C of a student, the lower is the outcome Z.	Detrimental Positive Misfit & Beneficial Negative Misfit Model: $Z = b_0 + b_1S + b_2C + \varepsilon$ with $b_1 < 0, b_2 > 0$
8. Accurate Self-Concept Hypothesis	For two students, the outcome Z is higher for the person whose self-concept S is closest to competence C.	Accurate Self-Concept Model: $Z = b_0 + b_1S + b_2C + b_3S^2 + b_4SC + b_5C^2 + \varepsilon$ with $b_1 = b_2 = 0, b_4 = -2b_3, b_5 = b_3, b_3 < 0$
9. Optimal Margin Hypothesis	For two students, the outcome Z is higher for the person whose positive discrepancy between self-concept S and competence C is closer to a constant K.	Optimal Margin Model: $Z = b_0 + b_1S + b_2C + b_3S^2 + b_4SC + b_5C^2 + \varepsilon$ with $b_2 = -b_1, b_4 = -2b_3, b_5 = b_3, b_3 < 0, K := (b_1-b_2)/(4*b_3)<0$
10. Curvilinear PSC Hypothesis	The association of self-concept S and the outcome Z diminishes at higher levels of self-concept or even becomes negative after some inflection point.	Curvilinear PSC Model: $Z = b_0 + b_1S + b_3S^2 + \varepsilon$ with $b_3 < 0$
11. Curvilinear Competence Hypothesis	The association of competence C and the outcome Z diminishes at higher levels of competence or even becomes negative after some inflection point.	Curvilinear Competence Model: $Z = b_0 + b_2C + b_5C^2 + \varepsilon$ with $b_5 < 0$
12. Interaction Hypothesis	The association of self-concept S and the outcome Z is more positive / less negative at higher levels of competence C than at lower levels of competence C.	Interaction Model: $Z = b_0 + b_1S + b_2C + b_4SC + \varepsilon$ with $b_4 > 0$
13. Global Hypothesis	Global Model	$Z = b_0 + b_1S + b_2C + b_3S^2 + b_4SC + b_5C^2 + \varepsilon$
00. Null Hypothesis	Null Model	$Z = b_0$

Note. Z denotes the outcome variable, S the self-concept, and C the competence.

for the predictors. Thus, we could investigate effects across a time span of at least 2 years. See Table 2 for an overview of the items, the waves in which the measures were administered, the internal consistencies (Cronbach's alphas) of the scales, and the reliabilities of the competence tests. As indicators, we used Weighted Likelihood Estimates (WLEs) for the competence tests and the scale means for all other variables.

Self-Concept

We used measures of students' academic self-concepts in math, reading, and school in general. First created for use in PISA 2000, three short scales for measuring academic self-concept were included in the NEPS in Wave 1 comparable to standard measures of academic self-concept (Wohlkinger

Table 2. Overview of Variables, Item Wording, and Sample Sizes in Each Wave

	Construct / Variable	Scale reference	Item wording	Cronbach's alpha / reliability of WLEs	Response format	Wave (Grade)	Year of data collection	Perspective	Wave sample size
Predictors									
	Reading competence	Competence test – WLE, linked (Fischer et al., 2016)	33 items (Pohl et al., 2012)	.77	Metric	1 (5)	Nov 2010 – Jan 2011	Test	5,208
	Math competence	Competence test – WLE, linked (Fischer et al., 2016)	25 items (Duchhardt & Gerdes, 2012)	.78	Metric	1 (5)	Nov 2010 – Jan 2011	Test	5,208
	Reading self-concept	NEPS (Wohlklinger et al., 2016)	How well do you read? a) I sometimes have trouble understanding a text really well. (reverse coded) b) I can understand texts very well and quickly. c) I have to read many things several times before I fully understand them. (reverse coded)	.82	4-point Likert scale	1 (5)	Nov 2010 – Jan 2011	Self	5,208
	Math self-concept	NEPS (Wohlklinger et al., 2016)	How would you rate your performance in school? a) I get good grades in math. b) Math is one of my best subjects. c) I have always been good at math.	.86	4-point Likert scale	1 (5)	Nov 2010 – Jan 2011	Self	5,208
	Academic self-concept	NEPS (Wohlklinger et al., 2016)	How would you rate your performance at school? a) I learn fast in most of the school subjects. b) In most of the school subjects, I perform well in written class tests. c) I perform well in most of the school subjects.	.82	4-point Likert scale	1 (5)	Nov 2010 – Jan 2011	Self	5,208
Outcomes									
Later educational achievement	Reading competence	Competence test – WLE, linked (Fischer et al., 2016)	42 items (Krannich et al., 2017)	.79		3 (7)	Nov 2012 – Jan 2013	Test	6,211
	Math competence	Competence Test – WLE, linked (Fischer et al., 2016)	23 items (Schnittjer & Gerken, 2017)	.72		3 (7)	Nov 2012 – Jan 2013	Test	6,211
Well-being	Stress	Gross &	The following part is about your personal situation in	.70	5-point	4 (8)	Nov	Self	5,558

Construct / Variable	Scale reference	Item wording	Cronbach's alpha / reliability of WLEs	Response format	Wave (Grade)	Year of data collection	Perspective	Wave sample size
	Seebaß (2014): The Standard Stress Scale	general. Please consider all areas of your life. To what extent do the following statements apply to you? a) If I don't enjoy a certain activity, then I usually don't have to do it. b) If I don't take care of something myself, nobody else will. (reverse coded) c) I pursue useful activities. d) I often feel lonely. (reverse coded) e) My achievements are suitably appreciated. f) There are people on whom I can rely. g) I generally get a good night's sleep. h) I think about problems a lot. (reverse coded) i) I feel exhausted after a normal day. (reverse coded) j) I worry about what my life might look like in three years. (reverse coded) k) I'm looking forward to the future.		Likert scale		2013 – May 2014		
Self-esteem	Collani & Herzberg (2003)	To what extent do the following statements apply to you? a) All in all, I am satisfied with myself. b) Now and then, I think that I'm not good for anything. (reverse coded) c) I have some positive attributes. d) I can do many things just as well as most other people. e) I am afraid there is not much I can be proud of. (reverse coded) f) Sometimes I really feel useless. (reverse coded) g) I consider myself a valuable person, at least I am not less valuable than the others. h) I wish I could have more respect for myself. (reverse coded) i) All in all, I tend to consider myself a loser. (reverse coded) j) I have a positive attitude towards myself.	.90	5-point Likert scale	5 (9)	Nov 2014 – Jan 2015	Self	4,898
Satisfaction	Following Cummins & Lau (2005) Personal Well-being Index: School Children (item	How satisfied are you … a) … with your life overall at the present? b) … with what you have? Think of money and things that you own. c) … with your health? d) … your family life? e) … your group of friends and acquaintances?	.82	10-point Likert scale	4 (8)	Nov 2013 – May 2014	Self	5,558

	Construct / Variable	Scale reference	Item wording	Cronbach's alpha / reliability of WLEs	Response format	Wave (Grade)	Year of data collection	Perspective	Wave sample size
		f added by NEPS)	f) … with your situation at school?						
Social relationships	Self-reported strengths and difficulties: peer problem scale	Goodman (1997): Strength and Difficulties Questionnaire	Please give a description of yourself. Think of the last half year! a) Most of the time I spend alone; I rather concentrate on myself. (reverse coded) b) I have one or several good friends. c) Generally, I am popular with other peers. d) I am teased or harassed by others. (reverse coded) e) I get along better with adults than with peers. (reverse coded)	.55	3-point Likert scale	6 (9)	Mar 2015 – July 2015	Self	4,627
	Parental reports on strengths and difficulties: peer problem scale	Goodman (1997): Strength and Difficulties Questionnaire	Now I would like to ask some more questions about your views on <name of target child>. This time I will state some characteristics. I would like you to assess whether they apply to <name of target child>. Please consider the behavior of <name of target child> during the previous six months for you answer. a) Mostly plays alone. (reverse coded) b) Has at least one good friend. c) Generally popular with other children. d) Is teased or victimized by others. (reverse coded) e) Gets along better with adults than other children. (reverse coded)	.59	3-point Likert scale	6 (9)	Mar 2015 – July 2015	Parent	3,262
	Number of friends	NEPS	Now let's talk about <name of target child>'s friends. How many different friends does <name of target child> meet with regularly in his/her free time? If you are not completely sure, please estimate the number.		metric	3 (7)	Nov 2012 – Jan 2013	Parent	4,638
	Social integration	NEPS (Dahm et al., 2016)	Now I would like to ask you some questions about <name of target child>'s school day. To what extent do the following statements apply to <name of target child>? a) <name of target child> has become well-integrated in class. b) <name of target child> is friends with many of the children in the class. c) <target child's name> has made new friends in class.	.61	4-point Likert scale	4 (8)	Nov 2013 – May 2014	Parent	4,187

Note. "NEPS" indicates that the scale was developed specifically by the NEPS administration. Other scales are listed with a reference or a note about further development by the NEPS administration. The Cronbach's alpha score is based on the data used in this study, the reliability of the WLEs was computed and provided in the respective technical report.

et al., 2016). For detailed descriptions of the item wordings and scale references, see Table 2.

Competence and Educational Achievement

We used students' performance on standardized competence tests in math and reading as indicators of competence as predictors and as educational achievement outcome measures. These tests were low-stakes tests administered by the NEPS in Waves 1 and 3; that is, results were only relevant in the context of the NEPS and had no consequences for students' educational trajectories. The tests for math competence were developed in accordance with German educational standards on the one hand (KMK, 2004, 2005)and the literacy framework of the *Programme for International Student Assessment* (PISA) (Weinert et al., 2011). By combining the two frameworks, NEPS math competence tests can be regarded as sound measures of educational success as well as competence in certain domains of adult life (Neumann et al., 2013).

The math test consisted of 25 items in Wave 1 (Grade 5) and 23 items in Wave 3 (Grade 7). The test targets students' mathematical competence (Weinert et al., 2011) in the sense of an individual's ability to comprehend the role of mathematics in everyday life as well as the ability to make well-considered mathematical judgments (Organisation for Economic Co-Operation and Development, 2003). The theoretical framework of the NEPS differentiates between mathematical content areas and the cognitive processes that are necessary to solve the tasks. In the content areas "quantity," "change and relationships," "space and shape," and "data and chance," students are asked to use cognitive processes, such as exerting technical skills, modeling, and problem solving (Ehmke et al., 2009; Weinert et al., 2011). The four content areas are assessed in equal shares, with the cognitive components distributed across the items (Schnittjer & Duchhardt, 2015). In Grade 5, students had to answer single-choice questions about tasks that involved spatial thinking, such as calculating surfaces For example items from the students' tests in Grades 5 to 7, please see Schnittjer and Duchhardt (2015).

The framework for reading competence in the NEPS is strongly related to the OECD´s emphasis on relevance in everyday life (Artelt et al., 2013). Therefore, the measurement of reading competence in the NEPS is expected to provide insight into educational processes, to be comparable to international assessments, and to be sufficient for meeting national standards (Artelt et al., 2013). The NEPS framework for reading competence targets students' abilities to retrieve information from common types of text that they could encounter in daily life, to integrate and interpret text fragments, and to reflect on and evaluate information from texts (Gehrer et al., 2012, 2013; Weinert et al., 2011). The tasks for reading competence are adjusted to students' age level in terms of difficulty and selected topics (Weinert et al., 2011). For example, in Grade 5, students had to read a short text on Egypt as an ancient civilization and then answer text-related multiple- or single-choice questions.For example items, please see Gehrer et al. (2012).

Math and reading tests were administered via text booklets. Half of the students receive a booklet where the math test is followed by the reading test. For the other half, the order is reversed. Students were randomly assigned to one of the two conditions (Duchhardt & Gerdes, 2012). All students receive the same items for mathematic competence and reading competence in the same order (no multimatrix design within each of the tests; Duchhardt & Gerdes, 2012; Krannich et al., 2017; Pohl et al., 2012; Schnittjer & Gerken, 2017).

The competence tests were constructed specifically for the NEPS, and the item development underwent several stages, including quantitative and qualitative preliminary and pilot studies (Weinert et al., 2011). Item selection by item response theory and optimization ensured good psychometric properties for the main studies (Pohl & Carstensen, 2012). Because the items were developed for various age groups through a combination of linking studies and specific anchor item designs, the competence tests allowed for a high-quality assessment of competence and competence development across the lifespan (Weinert et al., 2011).

As an indicator of students' performance, we used the Weighted Likelihood Estimates (WLEs) for each test provided by the NEPS (Pohl & Carstensen, 2012). To have a measure of students' general competence, we used the mean of the math and reading tests per student (the correlation of the WLEs of reading and math competence was $r = .64$, $p < .001$). As predictors, we used WLEs in math and reading as well as their mean from Wave 1. For later math and reading competence as an outcome, we used the WLEs 2 years after the initial assessment from each student in Wave 3 (Grade 7) on the same scale (Fischer et al., 2016).

Well-Being

Regarding well-being, we used three self-report scales. The *Standard Stress Scale* (Gross & Seebaß, 2014) measured students' perception of stress in different aspects. *Self-esteem* was measured with the German version of *the Rosenberg Self-Esteem Scale* (von Collani & Herzberg, 2003). *Satisfaction* with various aspects of life measured students' satisfaction with different aspects of their daily life. The five-item scale was originally developed by Cummins and Lau (2005), and in the NEPS, a sixth item covering satisfaction with school life was added.

Social Relationships

Self-reported peer problems were assessed with five items from the Strength and Difficulties Questionnaire (Goodman, 1997) that described relationships with and behavior toward peers. We also used *observed peer problems* from the parental data set. The students' parents were asked to assess peer problems (Goodman, 1997) using items with very similar wording to the students' perspective (for small deviations in wording between self-assessed peer problems and parental reports, see Table 2). Parents also reported the child's *number of friends*. On a three-item scale (Dahm et al., 2016), the parents reported their children's *social integration* in class.

Analytical Strategy

To test how self-concept and competence jointly predict achievement, well-being, and social relationships, we compared the hypotheses summarized by Humberg et al. (2019) to identify the best fitting hypotheses. We compared these hypotheses for three pairs of predictors (academic self-concept and competence in math, reading, and in general) for each of our nine outcome variables. For the analysis of all statistical models, we used Mplus Version 7 (Muthén & Muthén, 1998 - 2012). Data preparation and model reduction took place in R (R Core Team, 2020). To generate the visualizations of the response surfaces, we used the RSA package (Schönbrodt & Humberg, 2021). All R-code and the output data from Mplus can be found in the Supplementary Material on the OSF (https://osf.io/m3px6/).

We tested the hypotheses on self-concept, competence, and their relationship with different combinations of predictors and outcome variables. As predictors, we always used one measure of competence (math, reading, general) and one measure of self-concept in the same domain. The nine outcomes were taken from the outcome domains of achievement, well-being, and social relationships. Each outcome was predicted by self-concept and competence in each of the three domains. The exception was academic achievement, where we did not predict math achievement from reading predictors, and we did not predict reading achievement from math predictors. Altogether, this resulted in 25 combinations of predictors and outcomes (see Table 3 for the full list). For each of the 25 predictor-outcome combinations, we tested the 14 hypotheses by specifying the respective statistical models as summarized by Humberg et al. (2019), resulting in a total of 350 specific models. For each statistical model, we calculated the second-order Akaike Information Criterion (AICc; Sugiura, 1978). Then for each of the 25 predictor-outcome combinations, we compared the AICc values of the 14 different hypotheses and selected the best fitting models for each predictor-outcome combination. Thus, the model with the best relative fit indicated support for the respective hypothesis.

To take the multilevel structure into account, we used the students' class membership as a cluster variable. Prior to all analyses, we z-standardized all predictors on the grand mean for commensurability (Humberg et al., 2018, 2019). As additional predictors in all models, we included the class-average of the predictor variables and centered the predictors on the class mean. We used the terms from the full polynomial model as the predictors on the class level (i.e., the class means for competence and self-concept, their quadratic terms, and their interaction term) to specify the models correctly. By using the full polynomial model on the class level across hypotheses, only the individual level varied between hypotheses (i.e., only the individual level contributed to variation in model fit across hypotheses). Missing data were treated by using the model-based full information maximum likelihood (FIML) approach (Enders, 2010). We specified all predictors as correlated on the individual level or the class level, respectively, to improve the handling of missing values and to reduce bias due to missing data (Graham, 2003).[2]

Next, we aimed to identify the best fitting hypotheses for each of the 25 outcome-predictor combinations. After estimating model fit regarding each hypothesis, we compared the fit within each set of the respective 14 statistical models for each of the 25 outcome-predictor combinations. Following prior research (Humberg et al., 2019), first, we reduced the set of models by removing the larger model of the two nested models because, in these cases, the larger model added only redundant free parameters to a simpler model (see Figure 1). In the model comparison within one predictor domain for one outcome, we excluded models with more parameters in favor of a model with fewer parameters. If their log-likelihood difference was smaller than or equal to 1, the model with more parameters was too similar to a directly nested model with fewer parameters (for a comparable strategy, see Humberg et al., 2019). After reducing the set of models, we computed Akaike weights for the remaining models using the second-order Akaike Information Criterion (Akaike, 1998; Burnham & Anderson, 2002; Sugiura, 1978). The Akaike weight of a model can be interpreted as the conditional probability that the model is the best-fitting model, given the data and the set of model candidates (Burnham & Anderson, 2002; Wagenmakers & Farrell, 2004). Furthermore, we included a 95% confidence set of models, including the model with the highest Akaike weight and all models with a cumulated Akaike weight exceeding 95% to ensure that the confidence set included the best model with a probability of at least 95% given the data and the specified models.

Results

In the following, we present the results from the model comparisons. For each of the three outcome categories and in each of the three domains, we tested which of the 14 hypotheses fit the data best. Tables 4 to 6 show the 95% confidence sets of models with the largest Akaike weights (which indicate the conditional probabilities that the respective models are the best-fitting models, added up to at least 95%). For each of the models in the confidence set, the tables present the Akaike weights, the estimated regression coefficients on the student level, and the amount of variance explained on the student level.

Information for all 350 models, including regression coefficients on the individual level, Log-Likelihood, AIC, Akaike weights, and information on model convergence, nested models, and results of model comparisons can be

2 This analytic procedure deviates from the preregistration, where we stated that we would reduce the bias due to missing data by specifying all predictor variables as correlated with each other and that we would include all predictor variables in all models as additional correlates. However, these models featured too many parameters and thus did not produce an estimation output, so we decided to specify the predictors as correlated within only one predictor domain (e.g., in a model of the mathematical predictor domain, only the parameters for math competence and math self-concept were correlated).

Table 3. List of 25 Predictor-Domain-Outcome Combinations

Predictor 1: competence domain (Wave 1)	Predictor 2: self-concept domain (Wave 1)	Outcome (Waves 3 to 6)
Educational achievement outcomes		
General	Academic	Math competence
Math	Math	Math competence
General	Academic	Reading competence
Reading	Reading	Reading competence
Well-being outcomes		
General	Academic	Stress
Math	Math	Stress
Reading	Reading	Stress
General	Academic	Self-esteem
Math	Math	Self-esteem
Reading	Reading	Self-esteem
General	Academic	Satisfaction
Math	Math	Satisfaction
Reading	Reading	Satisfaction
Social Relationship outcomes		
General	Academic	Self-report of peer problems
Math	Math	Self-report of peer problems
Reading	Reading	Self-report of peer problems
General	Academic	Parental report of peer problems
Math	Math	Parental report of peer problems
Reading	Reading	Parental report of peer problems
General	Academic	Number of friends
Math	Math	Number of friends
Reading	Reading	Number of friends
General	Academic	Social integration
Math	Math	Social integration
Reading	Reading	Social integration

found in the Supplementary Materials on the OSF (see the table "Results_for_all_models" at https://osf.io/m3px6/). Furthermore, we modeled the response surfaces that were included in the model confidence sets for all statistical models where the Global Hypothesis was specified. The response surfaces can likewise be found on the OSF (see Figures 1 to 3). Third, all Mplus model outputs are available on the OSF.

Results for the Educational Achievement Outcomes

For the educational achievement outcomes, the confidence sets mainly included more complex models, especially within the predictor domains of math and reading (see Table 4).

In both predictor domains regarding math and reading, the Global Model fit the data best (Akaike weight of w = 0.94 for later math competence, Akaike weight of w = 0.935 for later reading competence). That is, the data were best represented by the complex interplay of self-concept and competence, and neither of the other models could capture the empirical pattern alone. Both self-concept and competence had a positive effect on later achievement, and the effect of competence was more positive than the effect of self-concept. In the domain of math, the effect of self-concept was more pronounced at higher levels of competence. When math competence and math self-concept were both low, later math competence was the lowest (see Figure 2a). This interpretation of the response surface also applied to the very similar response surface of later reading competence (see Figure 2 b). Furthermore, the Interaction Model was also included in both cases, but the fit was comparatively worse. Moreover, the regression coefficients were very similar in these models.

In the predictor domain of general competence, the confidence sets included more models. For later math competence, the model that explained the data best was the Curvilinear Competence Model (w = 0.345), followed by the Global Model (w = 0.326). Furthermore, the set included the Beneficial PSC & Competence Model (w = 0.166) and the Beneficial Competence Model (w = 0.164): All models included in the confidence set suggested that for later math competence, general competence was the more important

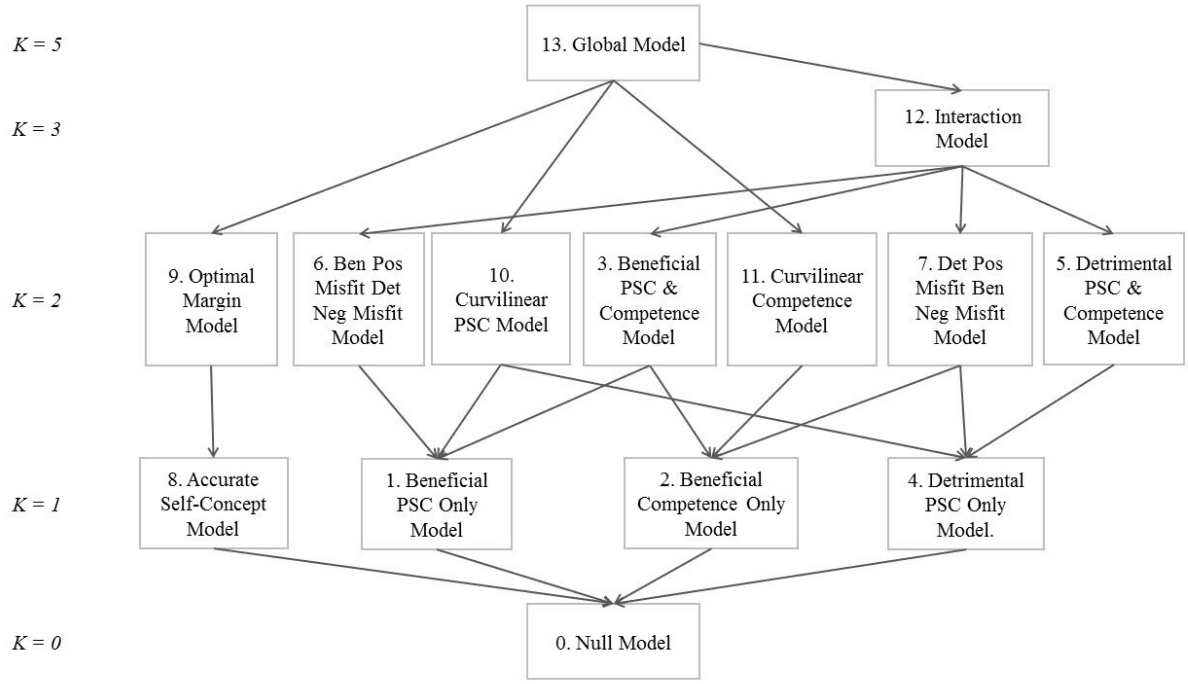


Figure 1. Nesting of the Models in the Initial Model Set

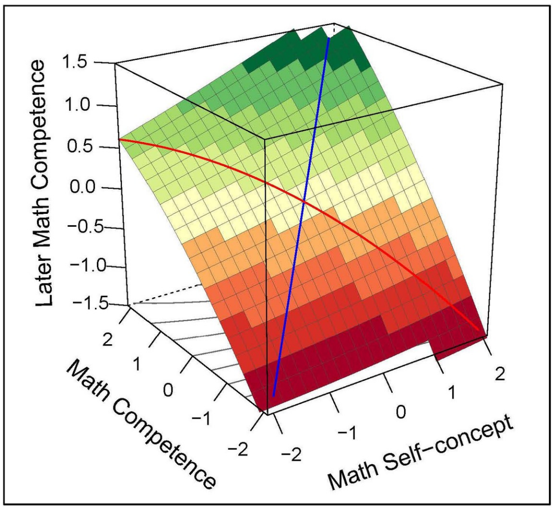
Note. A visualization of the nesting of the polynomial models in the initial model set. An arrow pointing form Model A to Model B indicates the nesting of Model B within Model A. K denotes the number of estimable parameters in relation to the Intercept-Only (Null) Model.

Table 4. Results of the Model Evaluation Analyses for the Educational Achievement Outcomes

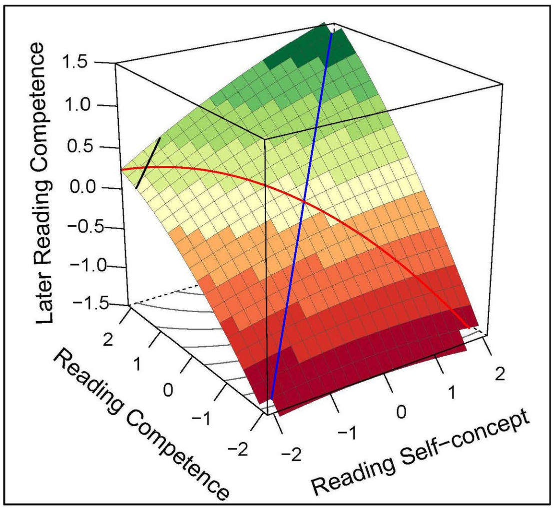
95% Confidence set of models	w	b_1	b_2	b_3	b_4	b_5	R^2_{within}
Outcome: Math Competence							
Domain of predictors: Math							
13. Global Model	0.944	0.16	0.669	0.007	0.057	-0.055	0.334
12. Interaction Model	0.03	0.158	0.656	0	0.03	0	0.330
Domain of predictors: General							
11. Curvilinear Competence Model	0.345	0	0.711	0	0	-0.035	0.284
13. Global Model	0.326	0.027	0.709	0.024	-0.015	-0.034	0.286
03. Beneficial PSC & Competence Model	0.166	0.022	0.702	0	0	0	0.284
02. Beneficial Competence Only Model	0.164	0	0.705	0	0	0	0.283
Outcome: Reading Competence							
Domain of predictors: Reading							
13. Global Model	0.935	0.153	0.601	-0.023	0.08	-0.057	0.213
12. Interaction Model	0.039	0.158	0.587	0	0.044	0	0.210
Domain of predictors: General							
03. Beneficial PSC & Competence Model	0.937	0.052	0.717	0	0	0	0.215
02. Beneficial Competence Only Model	0.063	0	0.724	0	0	0	0.212

Note. For each analysis, the 95% confidence set of models is provided. w = Akaike weight of the respective model = the model´s likelihood of being the best model in the set. Regression coefficients b_1 to b_5 refer to the full polynomial model $Z = b_0 + b_1S + b_2C + b3S^2 + b_4SC + b_5C^2$, where S is Self-Concept and C is Competence. R^2_{within} describes the variance explained on the individual level. The numbers before the model names refer to the list of models in Table 1.

predictor compared with self-concept. The Curvilinear Competence Model (which had the highest Akaike weight) indicated that general competence had a positive effect on later math competence, but it was weaker at higher levels of general competence. For later reading competence, the Beneficial PSC & Competence Model explained the data



a. Global Model for later math competence in the predictor domain of math competence / self-concept



b. Global Model for later reading competence in the predictor domain of reading competence / self-concept

Figure 2. Response Surfaces for the Global Models Leading the Confidence Sets for the Educational Achievement Outcomes

Note. Response surfaces for Global Models that led the confidence sets for the educational achievement outcomes. Predictors are denoted on the horizontal axes (competence on the left, self-concept on the right), and outcomes are represented on the vertical axis.

best (w = 0.937), meaning that both general competence and self-concept had a positive effect on later reading competence.

Overall, educational achievement was explained best by models that were more complex than simple linear models. In all models included in the confidence sets, competence (math, reading, and general competence) was the strongest predictor of educational achievement. There was no evidence that the misfit of self-concept and competence predicted later educational achievement.

Results for Well-Being Outcomes

For the outcome of stress, across all three predictor domains, models indicating a beneficial effect of self-concept explained the data best (see Table 5.

In the predictor domain of math, stress was explained best by the Curvilinear PSC Model (w = 0.6) followed by the Beneficial PSC Only Model (w = 0.29) and the Global Model (w = 0.11). Therefore, according to the Curvilinear PSC Model, mathematical self-concept had a beneficial effect on stress that was weaker at higher levels of mathematical self-concept. In the predictor domain of reading, the Beneficial Positive Misfit & Detrimental Negative Misfit Model was the best-fitting model for explaining stress (w = 0.334), followed by the Curvilinear PSC Model (w = 0.3), which fit almost as well, and the Global Model (w = 0.21). Thus, according to the Beneficial Positive Misfit & Detrimental Negative Misfit Model, students who overestimated their reading competence more had lower levels of stress. The Curvilinear PSC Model, which fit the data equally well, showed a beneficial effect of reading self-concept on stress that was weaker at higher levels of self-concept. In the predictor domain of general competence, the Beneficial PSC Only Model best explained the data for stress (w = 0.799). That is, in this predictor domain, general self-concept was the sole predictor of stress. Overall, our results suggest that stress was best explained by a positive self-concept. A positive misfit of self-concept and competence could also be an explanation for stress, but it appeared less likely.

For the outcome of self-esteem, models including a positive self-concept with competence as an additional predictor explained the data best. Here, in both in the predictor domains of math (w = 0.865) and general competence (w = 0.988), the Beneficial PSC & Competence Model was by far the best-fitting model. The Beneficial PSC & Competence Model indicated that students with higher general or mathematical self-concept had higher self-esteem, even when we controlled for general or math competence. In the predictor domain of reading, the Beneficial PSC Only Model (w = 0.649) fit the data best, meaning that reading self-concept was the sole predictor of self-esteem.

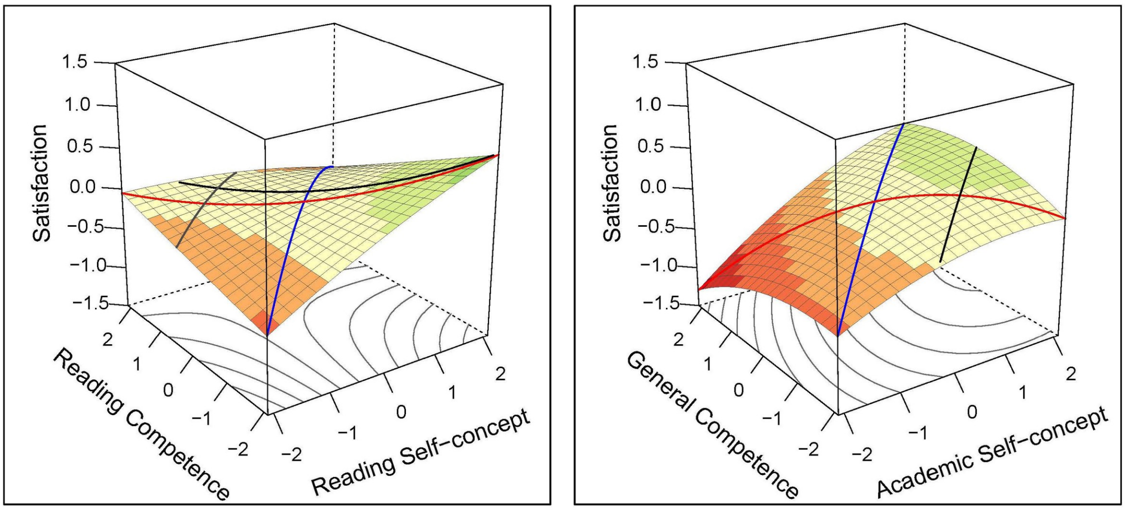
For the outcome of life satisfaction, once again more complex models were the best-fitting models in the predictor domains of both reading and general competence. That is, all confidence sets included the Global, the Curvilinear PSC, and the Beneficial Positive Misfit & Detrimental Negative Misfit Models. For the predictor domain of math, the Curvilinear PSC Model fit the data best in the model set (w = 0.73). According to this model, students with higher mathematical self-concept had a higher level of life satisfaction. However, this beneficial effect flattened at higher levels of mathematical self-concept. In the reading (w = 0.68) and general competence domains (w = 0.528), the best model for

Table 5. Results of the Model Evaluation Analyses for the Well-Being Outcomes

95% Confidence set of models	W	b_1	b_2	b_3	b_4	b_5	R^2_{within}
Outcome: Stress							
Domain of predictors: Math							
10. Curvilinear PSC Model	0.6	0.063	0	-0.016	0	0	0.016
01. Beneficial PSC Only Model	0.289	0.068	0	0	0	0	0.015
13. Global Model	0.111	0.064	-0.005	-0.019	0.019	-0.013	0.016
Domain of predictors: Reading							
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.334	0.048	-0.021	0	0	0	0.007
10. Curvilinear PSC Model	0.297	0.039	0	-0.016	0	0	0.006
13. Global Model	0.21	0.044	-0.02	-0.013	-0.015	0	0.008
01. Beneficial PSC Only Model	0.132	0.043	0	0	0	0	0.006
Domain of predictors: General							
01. Beneficial PSC Only Model	0.799	0.077	0	0	0	0	0.019
13. Global Model	0.201	0.076	-0.005	-0.007	0.024	-0.016	0.020
Outcome: Self-esteem							
Domain of predictors: Math							
03. Beneficial PSC & Competence Model	0.865	0.127	0.055	0	0	0	0.036
13. Global Model	0.123	0.123	0.056	-0.02	0.016	-0.007	0.037
Domain of predictors: Reading							
01. Beneficial PSC Only Model	0.649	0.062	0	0	0	0	0.006
12. Interaction Model	0.35	0.063	0.001	0	0.031	0	0.007
Domain of predictors: General							
03. Beneficial PSC & Competence Model	0.988	0.116	0.055	0	0	0	0.028
Outcome: Life satisfaction							
Domain of predictors: Math							
10. Curvilinear PSC Model	0.73	0.15	0	-0.073	0	0	0.014
13. Global Model	0.216	0.164	-0.053	-0.071	0.003	-0.02	0.015
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.035	0.186	-0.06	0	0	0	0.013
Domain of predictors: Reading							
13. Global Model	0.68	0.113	-0.067	-0.02	-0.092	0.002	0.008
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.213	0.121	-0.069	0	0	0	0.005
10. Curvilinear PSC Model	0.06	0.096	0	-0.039	0	0	0.005
Domain of predictors: General							
13. Global Model	0.528	0.243	-0.055	-0.046	0.054	-0.052	0.028
10. Curvilinear PSC Model	0.27	0.236	0	-0.042	0	0	0.027
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.132	0.253	-0.059	0	0	0	0.025
01. Beneficial PSC Only Model	0.07	0.247	0	0	0	0	0.024

Note. For each analysis, the 95% confidence set of models is provided. w = Akaike weight of the respective model = the model´s likelihood of being the best model in the set. Regression coefficients b_1 to b_5 refer to the full polynomial model $Z = b_0 + b_1S + b_2C + b3S^2 + b_4SC + b_5C^2$, where S is Self-Concept and C is Competence. R^2_{within} describes the variance explained on the individual level of the multilevel models. The numbers before the model names refer to the list of models in Table 1.

explaining life satisfaction was the Global Model, meaning that neither of the other models could capture the empirical pattern alone. The response surface for the effects of reading competence and reading self-concept on life satisfaction showed that the beneficial effect of reading self-concept decreased with higher levels of reading competence. Life satisfaction was lowest with low reading competence and low self-concept and was highest for people with a misfit between self-concept and competence, in which reading self-concept was more positive than reading competence (see Figure 3a). The response surface for life satisfaction in the general predictor domain was slightly different (see Figure



a. Global Model for Satisfaction in the predictor domain of reading competence / self-concept

b. Global Model for Satisfaction in the predictor domain of general competence / academic self-concept

Figure 3. Response Surfaces for the Global Models Leading the Confidence Sets of the Well-Being Outcomes

Note. Response surfaces for Global Models that led the confidence sets for the well-being outcomes. Predictors are denoted on the horizontal axes (competence on the left, self-concept on the right), and outcomes are represented on the vertical axis.

3b). Students had the lowest life satisfaction when they had high general competence but low academic self-concept, and they had the highest life satisfaction with higher academic self-concept and higher general competence. Here, self-concept played a larger role than general competence. All confidence sets with satisfaction as the outcome included models indicating positive linear or curvilinear effects of self-concept on satisfaction.

Overall, for well-being as the outcome, a positive self-concept was more important than competence or the fit or misfit between self-concept and competence.

Results for Social Relationship Outcomes

For social relationship outcomes, two main differences emerged in comparison with the achievement and well-being outcomes. First, more diverse and less consistent models were included in the confidence sets, for example, the confidence set for the outcome of parental reports of peer problems in the predictor domain of math included contradictory models, such as the Accurate Perception Model and the Beneficial Positive Misfit & Detrimental Negative Misfit Model (see Table 6). Second, the amount of variance explained on the student level was comparatively low. Both indicated that all models with social relationships as the outcome did not fit the data well. Thus, we interpret the results only briefly and advise caution when interpreting these models. For math and reading competence as predictor domains, across all outcomes, similar results emerged. Overall, the explained variance was low, and the confidence sets included several and diverse models. However, they had in common a pronounced negative effect of competence that was comparatively stronger than the other effects. Self-concept and its misfit with competence did not seem to be relevant predictors in these models.

For general competence as the predictor domain, across all social relationship outcomes, more complex models explained the data best, and the confidence sets were often led by the Optimal Margin Model or the Global Model. Across most outcomes, the effects of self-concept were equally as strong as the effects of competence; however, competence still had negative effects on all outcomes. Overall, for the social relationship outcomes, the confidence sets were larger than in other outcome domains. Between outcomes as well as between predictor domains, the models that were included in the confidence sets varied greatly. In sum, the results were less conclusive than in the achievement and well-being outcome domains.

Discussion

The present study tested competing hypotheses on the effects of competence and self-concept on educational achievement, well-being, and social relationships 2 to 5 years later as well as whether participants' self-perceptions were in line with their competence (showed fit) or not. We investigated three different competence domains (math, reading, and general) and nine outcomes. The sample (N = 6,086) we analyzed came from a large panel study (NEPS) on students in Grades 5 (Wave 1) to 9 (Waves 5 and 6). For each combination of predictor domain and outcome, the competing hypotheses were compared with regard to their model fit. The results differed across outcome domains. Achievement was best explained by a complex configuration of competence and self-concept. Roughly speaking, achievement was more positive the more positive

Table 6. Results of the Model Evaluation Analyses for the Social Relationship Outcomes

95% Confidence set of models	W	b_1	b_2	b_3	b_4	b_5	R^2_{within}
Outcome: Self-assessed peer problems							
Domain of predictors: Math							
10. Curvilinear PSC Model	0.264	0.001	0	-0.014	0	0	0.002
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.253	0.009	-0.017	0	0	0	0.001
11. Curvilinear Competence Model	0.167	0	-0.012	0	0	-0.006	0.001
00. Null Model	0.165	0	0	0	0	0	
13. Global Model	0.15	0.005	-0.015	-0.013	0.001	-0.005	0.003
Domain of predictors: Reading							
11. Curvilinear Competence Model	0.807	0	-0.014	0	0	-0.016	0.004
13. Global Model	0.127	0.001	-0.015	-0.009	0	-0.015	0.005
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.025	0.004	-0.02	0	0	0	0.002
Domain of predictors: General							
09. Optimal Margin Model	0.634	0.026	-0.026	-0.005	0.011	-0.005	0.008
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.216	0.027	-0.026	0	0	0	0.007
13. Global Model	0.147	0.027	-0.023	-0.002	0.008	-0.017	0.009
Outcome: Parent reports of peer problems							
Domain of predictors: Math							
13. Global Model	0.443	0.003	-0.025	-0.015	0.026	-0.017	0.008
08. Accurate Self-Concept Model	0.202	0	0	-0.016	0.031	-0.016	0.005
11. Curvilinear Competence Model	0.177	0	-0.023	0	0	-0.01	0.005
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.103	0.007	-0.03	0	0	0	0.005
05. Detrimental PSC & Competence Model	0.069	0	-0.028	0	0	0	0.004
Domain of predictors: Reading							
05. Detrimental PSC & Competence Model	0.266	-0.005	-0.019	0	0	0	0.003
11. Curvilinear Competence Model	0.252	0	-0.019	0	0	-0.004	0.003
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.213	0	-0.021	0	0	0	0.003
13. Global Model	0.152	-0.007	-0.018	-0.007	-0.017	0	0.005
00. Null Model	0.063	0	0	0	0	0	
10. Curvilinear PSC Model	0.054	-0.013	0	-0.011	0	0	0.002
Domain of predictors: General							
13. Global Model	0.575	0.009	-0.028	-0.011	0.015	-0.017	0.009
11. Curvilinear Competence Model	0.199	0	-0.025	0	0	-0.015	0.006
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.146	0.013	-0.032	0	0	0	0.006
08. Accurate Self-perception	0.043	0	0	-0.012	0.024	-0.012	0.005

95% Confidence set of models	W	b_1	b_2	b_3	b_4	b_5	R^2_{within}
Model							
Outcome: Parent reports of number of friends							
Domain of predictors: Math							
09. Optimal Margin Model	0.464	0.149	-0.149	-0.04	0.081	-0.04	0.003
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.354	0.141	-0.189	0	0	0	0.003
01. Beneficial PSC Only Model	0.055	0.094	0	0	0	0	0.001
11. Curvilinear Competence Model	0.05	0	-0.115	0	0	-0.065	0.001
00. Null Model	0.046	0	0	0	0	0	
Domain of predictors: Reading							
11. Curvilinear Competence Model	0.33	0	-0.189	0	0	-0.023	0.003
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.303	0.004	-0.198	0	0	0	0.003
05. Detrimental PSC & Competence Model	0.303	0	-0.197	0	0	0	0.003
00. Null Model	0.041	0	0	0	0	0	
Domain of predictors: General							
11. Curvilinear Competence Model	0.604	0	-0.178	0	0	-0.129	0.003
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.175	0.045	-0.214	0	0	0	0.002
05. Detrimental PSC & Competence Model	0.135	0	-0.206	0	0	0	0.002
09. Optimal Margin Model	0.056	0.098	-0.098	-0.036	0.071	-0.036	0.002
Outcome: Parent reports of social integration							
Domain of predictors: Math							
11. Curvilinear Competence Model	0.451	0	-0.039	0	0	-0.013	0.004
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.289	0.007	-0.047	0	0	0	0.004
05. Detrimental PSC & Competence Model	0.246	0	-0.044	0	0	0	0.003
Domain of predictors: Reading							
11. Curvilinear Competence Model	0.493	0	-0.035	0	0	-0.021	0.005
13. Global Model	0.261	0.016	-0.041	-0.013	-0.016	-0.017	0.007
06. Beneficial Positive Misfit Detrimental Negative Misfit Model	0.186	0.019	-0.049	0	0	0	0.005
05. Detrimental PSC & Competence Model	0.058	0	-0.044	0	0	0	0.004
Domain of predictors: General							
09. Optimal Margin Model	0.516	0.054	-0.054	-0.016	0.031	-0.016	0.015
13. Global Model	0.45	0.05	-0.054	-0.013	0.014	-0.041	0.016

Note. For each analysis, the 95% confidence set of models is provided. w = Akaike weight of the respective model = the model´s likelihood of being the best model in the set. Regression coefficients b_1 to b_5 refer to the full polynomial model $Z = b_0 + b_1S + b_2C + b3S^2 + b_4SC + b_5C^2$, where S is Self-Perception and C is Competence. R^2_{within} describes the variance explained on the individual level of the multilevel models. The numbers before the model names refer to the list of models in Table 1.

self-concept and competence were. Well-being was best predicted by a positive self-concept or by a self-concept that was more positive than actual competence. Finally, social relationships (e.g., a large number of friends or infrequent problems with peers) were related to low competence. However, when the social relationship variables were used as outcomes, neither model seemed to fit the data very well.

Evidence of Effects of the Fit or Misfit of Self-Concept and Competence

Overall, evidence of effects of a fit or misfit between self-concept and competence was weak. In the outcome categories of educational achievement, well-being, and social relationships, models that included an effect of fit or misfit between self-concept and competence did not fit the observed data as well as models that did not include an effect of fit or misfit. Instead, models that included an effect of competence or positive self-concept fit the data better. Table 7 summarizes how often each hypothesis was featured in the confidence sets of the 25 predictor-outcome combinations. Please note that the structure and presentation of results in Table 7 led to the loss of information on the rank of each hypothesis in the confidence set of hypotheses for each respective predictor-outcome combination. This information as well as details on the respective predictor-outcome combination that the hypothesis was part of in the respective confidence set can be found in Table 7 in the Supplementary Materials on the OSF (https://osf.io/m3px6/).

In the model sets with achievement as the outcome, the hypotheses that fit the data best included a strong positive effect of prior competence as well as a positive effect of academic self-concept. The shape of the response surfaces in models that included the full polynomial suggested a multiplicative effect of prior competence and self-concept. That is, the positive effect of prior competence was stronger at higher levels of self-concept. Overall, this pattern of results matches previous educational achievement research (Paschke et al., 2020) that found that both competence and self-concept had a positive effect on later achievement. More generally, the results are in line with academic self-concept research that has repeatedly shown the added positive effects of self-concept and competence on later educational achievement (Arens et al., 2017; Marsh et al., 2018). Adding to self-concept research, we found that the full polynomial model might fit the data better than just the linear terms. That is, in addition to the linear effects of both self-concept and competence, we found that nonlinear and interaction effects also explained variance in later educational achievement.

Results for model sets with well-being as the outcome showed that hypotheses that included a positive effect of academic-self-concept on later well-being fit the data best. Hypotheses that favored an effect of a fit or misfit between self-concept and competence fit the data worse. These results are largely in line with the results on adults (Humberg et al., 2019). That is, the positive discrepancy or similarity between self-concept and the corresponding competence did not offer incremental explanatory value beyond the positive effects of self-concept and competence. This result further supports the importance of distinguishing between effects of a fit or misfit between self-concept and competence in addition to the effect of a positive self-concept. If we had tested our research questions by using only difference or residual scores instead of multiple components, our results would likely (and incorrectly) support a more pronounced effect of a fit or misfit. This notion has already been outlined conceptually and statistically before (Humberg et al., 2018; Kwan et al., 2004). Not distinguishing between fit and misfit on the one hand and a positive self-concept on the other hand might obfuscate actual systematic associations.

Results for social relationships were less straightforward. Here, several models fit the data well. However, the variance explained at the individual level was low. A major trend was that models that included a negative effect of competence fit the data better. That is, more competent students seemed to have a disadvantage in social relationships later on. This pattern is in line with studies that found negative social evaluations of students who displayed academic striving (Rentzsch et al., 2011, 2013). Finally, our findings for social relationship outcomes in a sample of adolescents differed from previous findings that showed beneficial effects of self-concept and competence on peer-rated agentic and communal outcomes in a sample of adults (Humberg et al., 2019).

Generalizing Evidence on the Effects of Self-Concept, Competence, and Their Fit or Misfit

Our results provided evidence of the beneficial effects of self-concept and competence, especially on educational achievement and well-being outcomes. Models that featured linear effects of self-concept and competence explained the data better than models that focused on the fit or misfit between self-concept and competence. However, our findings do not imply that the relationship between self-concept and competence does not matter because, for example, with educational achievement, complex hypotheses that included curvilinear and interaction effects of self-concept and competence fit the data even better than hypotheses that featured only linear effects of self-concept and competence.

When we compare our findings with previous results in adults (Humberg et al., 2019), we can conclude that effects on outcomes of educational achievement and well-being are similar between adults and adolescents, but outcomes involving social relationships differ. In detail, our result that competence predicted later educational achievement in adolescents is mirrored by studies using adult samples in which competence predicted job performance (e.g., Abas & Imam, 2016; Alsabbah & Ibrahim, 2017). Similarly, our result that self-concept predicted well-being better than competence did in adolescents was also found in adults (Humberg et al., 2019) and is in line with previous findings on students' personality characteristics, such as self-esteem and locus of control having a stronger effect on well-being than school grades (Huebner, 1991). Effects on social relationships apparently differ between adolescents and adults. For adolescents, we found negative effects of competence on all social relationship outcomes. During adolescence, achievement is sometimes evaluated negatively by students

Table 7. Summary of the Evidence in Favor of the Hypotheses: Frequencies of Featured Hypotheses in the Confidence Sets of Each Predictor-Outcome Combination

Hypothesis	Verbal description	Frequency of the hypothesis in the confidence sets of the respective predictor-outcome combinations		
		Educational achievement (four combinations)	Well-being (nine combinations)	Social relationships (12 combinations)
1. Beneficial Positive Self-Concept (PSC) Only Hypothesis	For two students, the outcome Z is higher for the person with the higher self-concept S.	0 of 4	5 of 9	1 of 12
2. Beneficial Competence Only Hypothesis	For two students, the outcome Z is higher for the person with the higher competence C.	2 of 4	0 of 9	0 of 12
3. Beneficial PSC & Competence Hypothesis	For two students with the same value (level) of competence C, the outcome Z is higher for the person with the higher self-concept S (and vice versa).	2 of 4	2 of 9	0 of 12
4. Detrimental PSC Only Hypothesis	For two students, the outcome Z is lower for the person with the higher self-concept S.	0 of 4	0 of 9	0 of 12
5. Detrimental PSC & Competence Hypothesis	For two students with the same value (level) of competence C, the outcome Z is lower for the person with the higher self-concept S (and vice versa).	0 of 4	0 of 9	6 of 12
6. Beneficial Positive Misfit & Detrimental Negative Misfit Hypothesis	The higher the positive discrepancy between self-concept S and competence C of a student, the higher is the outcome Z.	0 of 4	4of 9	11 of 12
7. Detrimental Positive Misfit & Beneficial Negative Misfit Hypothesis	The higher the positive discrepancy between self-concept S and competence C of a student, the lower is the outcome Z.	0 of 4	0 of 9	0 of 12
8. Accurate Self-Concept Hypothesis	For two students, the outcome Z is higher for the person whose self-concept S is closest to competence C.	0 of 4	0 of 9	1of 12
9. Optimal Margin Hypothesis	For two students, the outcome Z is higher for the person whose positive discrepancy between self-concept S and competence C is closer to a constant K.	0 of 4	0 of 9	4 of 12
10. Curvilinear PSC Hypothesis	The association of self-concept S and the outcome Z diminishes at higher levels of self-concept or even becomes negative after some inflection point.	0 of 4	5 of 9	2 of 12
11. Curvilinear Competence Hypothesis	The association of competence C and the outcome Z diminishes at higher levels of competence or even becomes negative after some inflection point.	1 of 4	0 of 9	10 of 12
12. Interaction Hypothesis	The association of self-concept S and the outcome Z is more positive / less negative at higher levels of competence C than at lower levels of competence C.	2 of 4	1 of 9	0 of 12
13. Global Hypothesis	Global Model	3 of 4	7 of 9	8 of 12
00. Null Hypothesis	Null Model	0 of 4	0 of 9	4 of 12

Note. The cells in the table indicate how often the given hypothesis was present in the 95% confidence set of the predictor-outcome combinations in an outcome category. Combinations refer to the number of predictor-outcome combinations in an outcome group and are theoretically the maximum number of frequencies for each cell in a column.

(e.g., Juvonen & Murdock, 1993; Pelkner et al., 2002). This effect can be due to behaviors, such as boasting (Juvonen & Murdock, 1993), or can be due to the perception that more competent students are less likable and more competitive (Rentzsch et al., 2011).

Methodologically, our study used the advantages offered by the information-theoretic approach: The information-theoretic approach helped to identify models that fit the data best in contrast to more traditional null hypothesis significance testing. In a traditional approach, every unlikely hypothesis would have been tested against a null hypothesis and could have been statistically significant even if its relative fit was suboptimal (Burnham & Anderson, 2002). Furthermore, our study benefitted from the theoretical precision of mathematical models (Humberg et al., 2019). Lastly, the use of confidence sets allowed us to interpret the results more broadly, especially when the Akaike weights of more than one or two models were very similar.

Finally, our study adds to the existing literature on effects of self-concept and competence on life outcomes (Humberg et al., 2019; Paschke et al., 2020; Schimmack & Kim, 2020) by using a large sample of adolescents and carefully developed standardized tests as the best possible measures of competence (Pohl & Carstensen, 2012; Warm, 1989). Additionally, our study adds to the literature because we controlled for context effects, such as the frame of reference (e.g., Bullock & Trombley, 1999; Marsh, 1987) in a multilevel model design. Furthermore, the longer prediction period between predictors and outcomes provided a broader picture of how effects of self-concept, effects of competence, and effects of accuracy (fit) or inaccuracy (misfit) between self-concept and competence unfold in the developmental stage of adolescence. All these methodological features support the robustness and generalizability of our findings as well as previous findings.

Limitations

The present study has some limitations that call for caution when interpreting the results. First, our study was not designed to draw conclusions about causality, at least not according to current standards in psychological research (see Grosz et al., 2020). When we refer to effects, we mean a theoretically based direction of movement from predictors to outcomes that we can substantiate only on the basis of the time lag between the predictors and outcomes.

Even though the present study allows predictive interpretations, the comparability of effects across outcome categories is somewhat impaired by the different time frames used for the outcomes. The educational achievement outcomes were assessed 1 to 2 years earlier than the well-being and social relationship outcomes. This time lag might in part explain the poorer model fit for social relationships. Furthermore, the effects across outcome categories were also difficult to compare because we were able to include competence as a predictor and as an outcome variable in the outcome category of educational achievement. The other two outcome categories did not have a corresponding covariate at Time 1. Future research could use data in which all outcomes are assessed at the same points in time but are assessed multiple times and could test whether the effects of the predictors on one outcome vary over time and whether the effects of the predictors across outcomes vary in magnitude or direction (see Singer & Willett, 2003, Chapter 5.3.).

Also, to draw conclusions for the German population of students, a study would need to use design weights to ensure representativeness (Gelman, 2007; Solon et al., 2015). However, deriving weights for all levels in a multilevel analysis is a complex endeavor in and of itself. To date, for the NEPS data, only weights at the student and school levels are available but (because of the respective sampling strategy) not at the class or course levels (Schnapp, 2020). Nevertheless, based on the large and systematically drawn sample, the coefficients were estimated with comparatively high precision.

Another limitation pertains to the commensurability of self-concept and competence. Commensurability of predictors means that the predictors are measured on the same scale and ideally represent the same content domain (e.g., Edwards, 2002). This would be the case for self-perception and peer perception measured on the same scale. To achieve commensurability, we z-standardized the predictors. On the one hand, this procedure might reduce the interpretability of the results. Although z-standardizing ensures that predictors that are measured on different scales (e.g., competence and self-concept) are now measured on the same interval scale, one has to interpret the results more cautiously. A one-unit change in z-standardized scales means a change of one standard deviation in the underlying scale. Yet, the underlying self-concept and competence scales might have different distributions and standard deviations (Edwards, 1994a; Edwards & Parry, 1993; Humberg et al., 2018).

We did not have the option to use measures of self-concept and competence with directly commensurable scales. To be more commensurable with the measure of self-concept, a sum score would have been more suitable for the competence measure. However, we chose to use weighted maximum likelihood estimates (WLEs; Warm, 1989) as measures of competence because, rather than using only the sum of the correct number of answers, these point estimates best represent the individual competence level that is most likely by considering only the observed responses and the assumed item response model while correcting for the bias in the maximum likelihood estimate (Pohl & Carstensen, 2012). Therefore, we chose a more precise measure of competence over a more commensurable measure.

We aimed to measure people's more general self-views of their own ability rather than to obtain a situation-specific assessment of a single performance. Self-concept, as it is defined by a frame of reference and the appraisals from, for example, the peer group (Bong & Skaalvik, 2003), comes closer to approximating the construct of self-perception that has been discussed in research on the effects of self-perception and its bias on external criteria (Colvin & Block, 1994; Taylor & Brown, 1988). Even though metacognitive measures might be more commensurable with competence tests as measures, the self-perception that is measured with metacognitive judgments expresses an individual's beliefs about their own competence with respect to a very specific task or item. However, domain-specific self-concept is not restricted to a single task or item but offers a broader and overarching retrospective account of one's perceived competence in a specific domain (Dapp & Roebers, 2021; Stankov & Crawford, 1997). Taken together, it is possible that neither measure is ideal in terms of commensurability. However, both measures represent the constructs of interest, namely, actual competence and self-concept in the most precise and valid way.

Conclusion

The present study offered insight into the effects of self-concept, competence, and their relationship in a large sample of students from Grades 5 to 9. In the school context, for educational achievement and well-being, a positive self-concept and actual competence seem more important than their fit or misfit. Students' later achievements and their well-being are higher when they are actually competent and evaluate themselves positively, but it does not matter whether they evaluate themselves in an accurate or biased way.

Contributions

Contributed to conception and design: SSc, TL, DS, AS
Contributed to acquisition of data: SSc, DS, AS
Contributed to analysis and interpretation of data: SSc, TL, DS, AS
Drafted and/or revised the article: SSc, TL, AS
Approved the submitted version for publication: SSc, TL, DS; AS

Acknowledgments

We thank Jane Zagorski for language editing.

Funding Information

One author (SSc) received internal funding from an equal opportunities grant from the equal opportunities office at the University of Bamberg. The funders had no role in the design of the study, data collection and analysis, decision to publish, or preparation of the manuscript. No additional external funding was received for this study.

Competing Interests

The authors declare that no competing interests exist.

Data Accessibility Statement

The data used in the present study were collected within the National Educational Panel Study (NEPS) in Germany. The data are publicly available for research, and thus, there are legal restrictions to our sharing of the data.

Initiated and funded by the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung, BMBF), the National Educational Panel Study (NEPS) is carried out by the Leibniz Institute for Educational Trajectories (LIfBi) at the University of Bamberg. It is run by an interdisciplinary network led by Prof. Dr. Cordula Artelt. Since 2014, the NEPS has been an infrastructure facility permanently financed under the funding agreement between the German Federal and State Governments: https://www.neps-data.de/Project-Overview. Thus, the data are not owned by the authors of the present study but by a third party. Access to the data is restricted due to data protection regulations and requires a data use agreement with LIfBi: https://www.neps-data.de/Data-Center/Data-Access/Data-Use-Agreements. All NEPS data sets and materials can be downloaded from the NEPS website by scientists who meet the criteria for access: https://www.neps-data.de/Data-Center/Data-and-Documentation.

Please contact the Research Data Center at LIfBi for any requests concerning the availability of NEPS data: https://www.neps-data.de/Data-Center/Contact-Data-Center. The authors did not have any extra privileges in accessing the data.

All R codes and the Output data from Mplus can be found in the Supplementary Material on the OSF (https://osf.io/m3px6/?view_only=1c25d68d4fd84492809c11a9272f4310)

Submitted: December 29, 2021 PDT, Accepted: July 14, 2022 PDT

References

Abas, M. C., & Imam, O. A. (2016). Graduates' Competence on Employability Skills and Job Performance. *International Journal of Evaluation and Research in Education*, *5*(2), 119–125. https://doi.org/10.11591/ijere.v5i2.4530

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Springer Series in Statistics, Perspectives in Statistics. Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer. https://doi.org/10.1007/978-1-4612-1694-0_15

Allen, J. P., Uchino, B. N., & Hafen, C. A. (2015). Running With the Pack: Teen Peer-Relationship Qualities as Predictors of Adult Physical Health. *Psychological Science*, *26*(10), 1574–1583. https://doi.org/10.1177/0956797615594118

Alsabbah, M. Y. A., & Ibrahim, H. I. (2017). Employees' Competence and Job Performance. *Advanced Science Letters*, *23*(1), 313–316. https://doi.org/10.1166/asl.2017.7170

Anderson, C., Brion, S., Moore, D. A., & Kennedy, J. A. (2012). A status-enhancement account of overconfidence. *Journal of Personality and Social Psychology*, *103*(4), 718–735. https://doi.org/10.1037/a0029395

Arens, A. K., Marsh, H. W., Pekrun, R., Lichtenfeld, S., Murayama, K., & vom Hofe, R. (2017). Math self-concept, grades, and achievement test scores: Long-term reciprocal effects across five waves and three achievement tracks. *Journal of Educational Psychology*, *109*(5), 621–634. https://doi.org/10.1037/edu0000163

Arens, A. K., 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*(4), 970–981. https://doi.org/10.1037/a0025047

Artelt, C., Weinert, S., & Carstensen, C. H. (2013). Assessing competencies across the lifespan within the German National Educational Panel Study (NEPS) - Editorial. *Journal for Educational Research Online*, *5*(2), 5–14.

Asendorpf, J. B., & Ostendorf, F. (1998). Is self-enhancement healthy? Conceptual, psychometric, and empirical analysis. *Journal of Personality and Social Psychology*, *74*(4), 955–966. https://doi.org/10.1037/0022-3514.74.4.955

Backé, E.-M., Seidler, A., Latza, U., Rossnagel, K., & Schumann, B. (2012). The role of psychosocial stress at work for the development of cardiovascular diseases: A systematic review. *International Archives of Occupational and Environmental Health*, *85*(1), 67–79. https://doi.org/10.1007/s00420-011-0643-6

Barranti, M., Carlson, E. N., & Côté, S. (2017). How to Test Questions About Similarity in Personality and Social Psychology Research. *Social Psychological and Personality Science*, *8*(4), 465–475. https://doi.org/10.1177/1948550617698204

Baumeister, R. F. (1989). The Optimal Margin of Illusion. *Journal of Social and Clinical Psychology*, *8*(2), 176–189. https://doi.org/10.1521/jscp.1989.8.2.176

Baumeister, R. F. (1999). Self-concept, self-esteem, and identity. In V. J. Derlega, B. A. Winstead, & W. H. Jones (Eds.), *Personality: Contemporary theory and research* (pp. 339–375). Nelson-Hall Publishers.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does High Self-Esteem Cause Better Performance, Interpersonal Success, Happiness, or Healthier Lifestyles? *Psychological Science in the Public Interest*, *4*(1), 1–44. https://doi.org/10.1111/1529-1006.01431

Bellmore, A. (2011). Peer rejection and unpopularity: Associations with GPAs across the transition to middle school. *Journal of Educational Psychology*, *103*(2), 282–295. https://doi.org/10.1037/a0023312

Birch, S. H., & Ladd, G. W. (1996). Interpersonal relationships in the school environment and children's early school adjustment: The role of teachers and peers. In J. Juvonen & K. R. Wentzel (Eds.), *Cambridge studies in social and emotional development. Social motivation: Understanding children's school adjustment* (pp. 199–225). Cambridge University Press. https://doi.org/10.1017/cbo9780511571190.011

Blossfeld, H.-P., & von Maurice, J. (2011). *Education as a lifelong process* (Vol. 14). Springer; VS-Verlag. https://doi.org/10.1007/s11618-011-0179-2

Bonanno, G. A., Rennicke, C., & Dekel, S. (2005). Self-enhancement among high-exposure survivors of the September 11th terrorist attack: Resilience or social maladjustment? *Journal of Personality and Social Psychology*, *88*(6), 984–998. https://doi.org/10.1037/0022-3514.88.6.984

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review*, *15*(1), 1–40. https://doi.org/10.1023/a:1021302408382

Box, G. E. P., & Draper, N. R. (1987). *Empirical model-building and response surfaces [7 Dr]*. Wiley.

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A Century of Grading Research. *Review of Educational Research*, *86*(4), 803–848. https://doi.org/10.3102/0034654316672069

Brown, J. D. (1986). Evaluations of Self and Others: Self-Enhancement Biases in Social Judgments. *Social Cognition*, *4*(4), 353–376. https://doi.org/10.1521/soco.1986.4.4.353

Bukowski, W. M., & Hoza, B. (1989). Popularity and friendship: Issues in theory, measurement, and outcome. In T. J. Berndt & G. W. Ladd (Eds.), *Wiley series on personlity processes. Peer relationships in child development* (pp. 15–45). Wiley.

Bullock, A., & Trombley, S. (1999). *The Norton Dictionary of modern thought* (2. rev. ed.). Norton.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2. ed.). Springer.

Casillas, A., Robbins, S., Allen, J., Kuo, Y.-L., Hanson, M. A., & Schmeiser, C. (2012). Predicting early academic failure in high school from prior academic achievement, psychosocial characteristics, and behavior. *Journal of Educational Psychology*, *104*(2), 407–420. https://doi.org/10.1037/a0027180

Chiu, M. M., & Klassen, R. M. (2009). Calibration of reading self-concept and reading achievement among 15-year-olds: Cultural differences in 34 countries. *Learning and Individual Differences*, *19*(3), 372–386. https://doi.org/10.1016/j.lindif.2008.10.004

Chiu, M. M., & Klassen, R. M. (2010). Relations of mathematics self-concept and its calibration with mathematics achievement: Cultural differences among fifteen-year-olds in 34 countries. *Learning and Instruction*, *20*(1), 2–17. https://doi.org/10.1016/j.learninstruc.2008.11.002

Chung, J., Schriber, R. A., & Robins, R. W. (2016). Positive Illusions in the Academic Context: A Longitudinal Study of Academic Self-Enhancement in College. *Personality & Social Psychology Bulletin*, *42*(10), 1384–1401. https://doi.org/10.1177/0146167216662866

Çivitci, N., & Çivitci, A. (2009). Self-esteem as mediator and moderator of the relationship between loneliness and life satisfaction in adolescents. *Personality and Individual Differences*, *47*(8), 954–958. https://doi.org/10.1016/j.paid.2009.07.022

Colvin, C. R., & Block, J. (1994). Do positive illusions foster mental health? An examination of the Taylor and Brown formulation. *Psychological Bulletin*, *116*(1), 3–20. https://doi.org/10.1037/0033-2909.116.1.3

Colvin, C. R., Block, J., & Funder, D. C. (1995). Overly positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology*, *68*(6), 1152–1162. https://doi.org/10.1037/0022-3514.68.6.1152

Cummins, R.A., & Lau, A. (2005). *Personal Wellbeing Index – School Children (PWI-SC*. Deakin University. http://www.acqol.com.au/uploads/pwi-sc/pwi-sc-english.pdf

Cummins, Robert A. (2005). The Domains of Life Satisfaction: An Attempt to Order Chaos. In *Citation Classics from Social Indicators Research* (pp. 559–584). Springer. https://doi.org/10.1007/1-4020-3742-2_19

Dahm, G., Lauterbach, O., & Hahn, S. (2016). Measuring Students' Social and Academic Integration—Assessment of the Operationalization in the National Educational Panel Study. In H.-P Blossfeld, J. von Maurice, M. Bayer, & J. Skopek (Eds.), *Methodological issues of longitudinal surveys: The example of the National Educational Panel Study* (pp. 313–329). Springer VS. https://doi.org/10.1007/978-3-658-11994-2_18

Dapp, L. C., & Roebers, C. M. (2021). Metacognition and self-concept: Elaborating on a construct relation in first-grade children. *PLOS ONE*, *16*(4), e0250845. https://doi.org/10.1371/journal.pone.0250845

Deventer, J., Humberg, S., Lüdtke, O., Nagy, G., Retelsdorf, J., & Wagner, J. (2019). Testing Competing Hypotheses on the Interplay of Importance and Support of the Basic Psychological Needs at Work and Personality Development with Response Surface Analysis. *Collabra: Psychology*, *5*(1). https://doi.org/10.1525/collabra.214

Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*(3), 542–575. https://doi.org/10.1037/0033-2909.95.3.542

Diener, E., Lucas, R. E., & Oishi, S. (2018). Advances and Open Questions in the Science of Subjective Well-Being. *Collabra: Psychology*, *4*(1), 15. https://doi.org/10.1525/collabra.115

Duchhardt, C., & Gerdes, A. (2012). *NEPS Technical Report for Mathematics: Scaling Results of Starting Cohort 3 in Fifth Grade* [NEPS Working Paper No. 19]. Otto-Friedrich-Universität, Nationales Bildungspanel.

Dufner, M., Gebauer, J. E., Sedikides, C., & Denissen, J. J. A. (2019). Self-Enhancement and Psychological Adjustment: A Meta-Analytic Review. *Personality and Social Psychology Review*, *23*(1), 48–72. https://doi.org/10.1177/1088868318756467

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension. *Current Directions in Psychological Science*, *16*(4), 228–232. https://doi.org/10.1111/j.1467-8721.2007.00509.x

Edwards, J. R. (1994a). The Study of Congruence in Organizational Behavior Research: Critique and a Proposed Alternative. *Organizational Behavior and Human Decision Processes*, *58*(1), 51–100. https://doi.org/10.1006/obhd.1994.1029

Edwards, J. R. (1994b). Regression Analysis as an Alternative to Difference Scores. *Journal of Management*, *20*(3), 683–689. https://doi.org/10.1177/014920639402000311

Edwards, J. R. (2002). Alternatives to difference scores: Polynomial regression analysis and response surface methodology. In F. Drasgow & N. Schmitt (Eds.), *The Jossey-Bass business & management series. Measuring and analyzing behavior in organizations: Advances in measurement and data analysis: Foreword by Neal Schmitt* (pp. 350–400). Jossey-Bass.

Edwards, J. R., & Parry, M. E. (1993). On the Use of Polynomial Regression Equations As An Alternative to Difference Scores in Organizational Research. *Academy of Management Journal*, *36*(6), 1577–1613. https://doi.org/10.5465/256822

Ehmke, T., Duchhardt, C., Geiser, H., Grüßing, M., & Heinze, A. (2009). Kompetenzentwicklung über die Lebensspanne – Erhebung von mathematischer Kompetenz im Nationalen Bildungspanel. In A. Heinze & M. Grüßing (Eds.), *Mathematiklernen vom Kindergarten bis zum Studium: Kontinuität und Kohärenz als Herausforderung für den Mathematikunterricht* (pp. 313–327). Waxmann.

Enders, C. K. (2010). *Applied missing data analysis. Methodology in the social sciences*. Guilford Press.

Fischer, L., Rohm, T., Gnambs, T., & Carstensen, C. H. (2016). *Linking the data of the competence tests* [NEPS Survey Paper No. 1].

Förster, N., Humberg, S., Hebbecker, K., Back, M. D., & Souvignier, E. (2022). Should teachers be accurate or (overly) positive? A competitive test of teacher judgment effects on students' reading progress. *Learning and Instruction*, *77*, 101519. https://doi.org/10.1016/j.learninstruc.2021.101519

Gehrer, K., Zimmermann, S., & Artelt, C. (2013). NEPS framework for assessing reading competence and results from an adult pilot study. *Journal for Educational Research Online*, *5*(2), 50–79.

Gehrer, K., Zimmermann, S., Artelt, C., & Weinert, S. (2012). *The assessment of reading competence (including sample items for grade 5 and 9): Scientific Use File 2012, Version 1.0.0*. University of Bamberg, National Educational Panel Study.

Gelman, A. (2007). Struggles with Survey Weighting and Regression Modeling. *Statistical Science*, *22*(2). https://doi.org/10.1214/088342306000000691

Giordano, P. C. (2003). Relationships in Adolescence. *Annual Review of Sociology*, *29*(1), 257–281. https://doi.org/10.1146/annurev.soc.29.010202.100047

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Goorin, L., & Bonanno, G. A. (2009). Would You Buy a Used Car from a Self-enhancer? Social Benefits and Illusions in Trait Self-enhancement. *Self and Identity*, *8*(2–3), 162–175. https://doi.org/10.1080/15298860802501607

Graham, J. W. (2003). Adding Missing-Data-Relevant Variables to FIML-Based Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(1), 80–100. https://doi.org/10.1207/s15328007sem1001_4

Gross, C., & Seebaß, K. (2014). *The Standard Stress Scale (SSS): Measuring stress in the life course* [NEPS Working Paper No. 45]. Otto-Friedrich-Universität, Nationales Bildungspanel.

Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The Taboo Against Explicit Causal Inference in Nonexperimental Psychology. *Perspectives on Psychological Science*, *15*(5), 1243–1255. https://doi.org/10.1177/1745691620921521

Guo, J., Nagengast, B., Marsh, H. W., Kelava, A., Gaspard, H., Brandt, H., Cambria, J., Flunger, B., Dicke, A.-L., Häfner, I., Brisson, B., & Trautwein, U. (2016). Probing the Unique Contributions of Self-Concept, Task Values, and Their Interactions Using Multiple Value Facets and Multiple Academic Outcomes. *AERA Open*, *2*(1), 233285841562688. https://doi.org/10.1177/2332858415626884

Guo, J., Parker, P. D., Marsh, H. W., & Morin, A. J. S. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology*, *51*(8), 1163–1176. https://doi.org/10.1037/a0039440

Hansen, D. J., Christopher, J. S., & Nangle, D. W. (1992). Adolescent heterosocial interactions and dating. In V. B. van Hasselt & M. Hersen (Eds.), *Handbook of Social Development: A Lifespan Perspective* (pp. 371–394). Springer Science and Business Media LLC. https://doi.org/10.1007/978-1-4899-0694-6_15

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

Harter, S. (2006). The self. In N. Eisenberg & W. Damon (Eds.), *Handbook of child psychology* (6th ed., Vol. 3, pp. 505–570). Wiley.

Heckman, J. J., & Kautz, T. (2012). Hard evidence on soft skills. *Labour Economics*, *19*(4), 451–464. https://doi.org/10.1016/j.labeco.2012.05.014

Heine, S. J., & Lehman, D. R. (1995). Cultural variation in unrealistic optimism: Does the West feel more vulnerable than the East? *Journal of Personality and Social Psychology*, *68*(4), 595–607. https://doi.org/10.1037/0022-3514.68.4.595

Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, *106*(4), 766–794. https://doi.org/10.1037/0033-295x.106.4.766

Hightower, E. (1990). Adolescent interpersonal and familial precursors of positive mental health at midlife. *Journal of Youth and Adolescence*, *19*(3), 257–275. https://doi.org/10.1007/bf01537891

Huebner, E. S. (1991). Correlates of life satisfaction in children. *School Psychology Quarterly*, *6*(2), 103–111. https://doi.org/10.1037/h0088805

Humberg, S., Dufner, M., Schönbrodt, F. D., Geukes, K., Hutteman, R., Küfner, A. C. P., van Zalk, M. H. W., Denissen, J. J. A., Nestler, S., & Back, M. D. (2019). Is accurate, positive, or inflated self-perception most advantageous for psychological adjustment? A competitive test of key hypotheses. *Journal of Personality and Social Psychology*, *116*(5), 835–859. https://doi.org/10.1037/pspp0000204

Humberg, S., Dufner, M., Schönbrodt, F. D., Geukes, K., Hutteman, R., van Zalk, M. H. W., Denissen, J. J. A., Nestler, S., & Back, M. D. (2018). Enhanced versus simply positive: A new condition-based regression analysis to disentangle effects of self-enhancement from effects of positivity of self-view. *Journal of Personality and Social Psychology*, *114*(2), 303–322. https://doi.org/10.1037/pspp0000134

IEA Data Processing and Research Center. (2012). *Methodenbericht NEPS Startkohorte 3: Haupterhebung - Herbst / Winter 2010*. IEA Data Processing and Research Center. https://www.neps-data.de/Datenzentrum/Daten-und-Dokumentation/Startkohorte-Klasse-5/Dokumentation

Jackson, L. D., & Bracken, B. A. (1998). Relationship Between Students' Social Status and Global and Domain-Specific Self-Concepts. *Journal of School Psychology*, *36*(2), 233–246. https://doi.org/10.1016/s0022-4405(97)00052-6

James, W. (1890). *The principles of psychology: Vol. I*. Henry Holt and Co. https://doi.org/10.1037/10538-000

John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality and Social Psychology*, *66*(1), 206–219. https://doi.org/10.1037/0022-3514.66.1.206

Juvonen, Jaana, & Murdock, T. B. (1993). How to promote social approval: Effects of audience and achievement outcome on publicly communicated attributions. *Journal of Educational Psychology*, *85*(2), 365–376. https://doi.org/10.1037/0022-0663.85.2.365

Khuri, A. I., & Cornell, J. A. (1987). *Response surfaces: Designs and analysis* (Vol. 81). Dekker.

Kitayama, S., & Karasawa, M. (1997). Implicit Self-Esteem in Japan: Name Letters and Birthday Numbers. *Personality and Social Psychology Bulletin*, *23*(7), 736–742. https://doi.org/10.1177/0146167297237006

Kiuru, N., Aunola, K., Lerkkanen, M.-K., Pakarinen, E., Poskiparta, E., Ahonen, T., Poikkeus, A.-M., & Nurmi, J.-E. (2015). Positive teacher and peer relations combine to predict primary school students' academic skill development. *Developmental Psychology*, *51*(4), 434–446. https://doi.org/10.1037/a0038911

KMK. (2004). *Bildungsstandards im Fach Mathematik für den Mittleren Schulabschluss (Jahrgangsstufe 10)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2003/2003_12_04-Bildungsstandards-Mathe-Mittleren-SA.pdf

KMK. (2005). *Bildungsstandards im Fach Mathematik für den Primarbereich (Jahrgangsstufe 4)*. https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_10_15-Bildungsstandards-Mathe-Primar.pdf

Krannich, M., Jost, O., Rohm, T., Koller, I., Pohl, S., Haberkorn, K., Carstensen, C. H., Fischer, L., & Gnambs, T. (2017). NEPS Technical Report for Reading: Scaling Results of Starting Cohort 3 for Grade 7. *NEPS National Education Panel Study*. https://doi.org/10.5157/NEPS:SP14:2.0

Kwan, V. S. Y., John, O. P., Kenny, D. A., Bond, M. H., & Robins, R. W. (2004). Reconceptualizing individual differences in self-enhancement bias: An interpersonal approach. *Psychological Review*, *111*(1), 94–110. https://doi.org/10.1037/0033-295x.111.1.94

Lansford, J. E., Sherman, A. M., & Antonucci, T. C. (1998). Satisfaction with social networks: An examination of socioemotional selectivity theory across cohorts. *Psychology and Aging*, *13*(4), 544–552. https://doi.org/10.1037/0882-7974.13.4.544

Leary, M. R. (2007). Motivational and emotional aspects of the self. *Annual Review of Psychology*, *58*(1), 317–344. https://doi.org/10.1146/annurev.psych.58.110405.085658

Luong, G., Charles, S. T., & Fingerman, K. L. (2011). Better With Age: Social Relationships Across Adulthood. *Journal of Social and Personal Relationships*, *28*(1), 9–23. https://doi.org/10.1177/0265407510391362

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, *79*(3), 280–295. https://doi.org/10.1037/0022-0663.79.3.280

Marsh, H. W., & Craven, R. G. (2006). Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. *Perspectives on Psychological Science*, *1*(2), 133–163. https://doi.org/10.1111/j.1745-6916.2006.00010.x

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology*, *81*(1), 59–77. https://doi.org/10.1348/000709910x503501

Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology*, *54*(2), 263–280. https://doi.org/10.1037/dev0000393

Marsh, H. W., & Shavelson, R. (1985). Self-Concept: Its Multifaceted, Hierarchical Structure. *Educational Psychologist*, *20*(3), 107–123. https://doi.org/10.1207/s15326985ep2003_1

Messick, S. (1984). The Psychology of Educational Measurement. *ETS Research Report Series*, *1984*(1), i–55. https://doi.org/10.1002/j.2330-8516.1984.tb00046.x

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, *130*(5), 711–747. https://doi.org/10.1037/0033-2909.130.5.711

Mielke, I., Humberg, S., Leckelt, M., Geukes, K., & Back, M. D. (2021). Do Narcissists Self-Enhance? Disentangling the Associations Between Narcissism and Positive Versus Enhanced Self-Views Across Aspects of Narcissism, Content Domains, and Comparison Criteria. *Social Psychological and Personality Science*, *12*(5), 580–592. https://doi.org/10.1177/1948550620930543

Mota, S., Leckelt, M., Geukes, K., Nestler, S., Humberg, S., Schröder-Abé, M., Schmukle, S. C., & Back, M. D. (2019). A Comprehensive Examination of Narcissists' Self-Perceived and Actual Socioemotional Cognition Ability. *Collabra: Psychology*, *5*(1). https://doi.org/10.1525/collabra.174

Muthén, L. K., & Muthén, B. O. (1998). *Mplus User's Guide* (Seventh Edition). Muthén & Muthén.

Negru-Subtirica, O., & Pop, E. I. (2016). Longitudinal links between career adaptability and academic achievement in adolescence. *Journal of Vocational Behavior*, *93*, 163–170. https://doi.org/10.1016/j.jvb.2016.02.006

Neumann, I., Duchhardt, C., Grüßing, M., & Heinze, A. (2013). Modeling and assessing mathematical competence over the lifespan. *Journal for Educational Research Online*, *5*(2), 80–109.

Newland, M. C. (2019). An Information Theoretic Approach to Model Selection: A Tutorial with Monte Carlo Confirmation. *Perspectives on Behavior Science*, *42*(3), 583–616. https://doi.org/10.1007/s40614-019-00206-1

Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of Objective and Subjective Career Success: A Meta-analysis. *Personnel Psychology*, *58*(2), 367–408. https://doi.org/10.1111/j.1744-6570.2005.00515.x

Olczyk, M., Will, G., & Kristen, C. (2016). Immigrants in the NEPS: Identifying generation status and group of origin. *NEPS National Education Panel Study*. https://doi.org/10.5157/NEPS:SP04:1.0

O'Mara, E. M., Gaertner, L., Sedikides, C., Zhou, X., & Liu, Y. (2012). A longitudinal-experimental test of the panculturality of self-enhancement: Self-enhancement promotes psychological well-being both in the west and the east. *Journal of Research in Personality*, *46*(2), 157–163. https://doi.org/10.1016/j.jrp.2012.01.001

Organisation for Economic Co-Operation and Development. (2003). *The PISA 2003 assessment framework – mathematics, reading, science and problem solving knowledge and skills*. OECD.

Paschke, P., Weidinger, A. F., & Steinmayr, R. (2020). Separating the effects of self-evaluation bias and self-view on grades. *Learning and Individual Differences*, *83–84*, 101940. https://doi.org/10.1016/j.lindif.2020.101940

Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A mixed blessing? *Journal of Personality and Social Psychology*, *74*(5), 1197–1208. https://doi.org/10.1037/0022-3514.74.5.1197

Pelkner, A.-K., Günther, R., & Boehnke, K. (2002). Die Angst vor sozialer Ausgrenzung als leistungshemmender Faktor? Zum Stellenwert guter mathematischer Schulleistungen unter Gleichaltrigen: [The fear of social rejection as an achievement inhibiting factor? About the status of good mathematical school achievements among peers]. *Zeitschrift Für Pädagogik*, *45*(Suppl.), 326–340. https://www.pedocs.de/volltexte/2013/7862/pdf/ZfPaed_45.Beiheft.pdf#page=327

Pfeifer, J. H., & Berkman, E. T. (2018). The Development of Self and Identity in Adolescence: Neural Evidence and Implications for a Value-Based Choice Perspective on Motivated Behavior. *Child Development Perspectives*, *12*(3), 158–164. https://doi.org/10.1111/cdep.12279

Pinquart, M., & Ebeling, M. (2020). Parental Educational Expectations and Academic Achievement in Children and Adolescents—a Meta-analysis. *Educational Psychology Review*, *32*(2), 463–480. https://doi.org/10.1007/s10648-019-09506-z

Pohl, S., & Carstensen, C. H. (2012). *NEPS Technical Report – Scaling the Data of the Competence Tests* [NEPS Working Paper No. 14]. Otto-Friedrich-Universität, Nationales Bildungspanel.

Pohl, S., Haberkorn, K., Hardt, K., & Wiegand, E. (2012). *NEPS Technical Report for Reading: Scaling Results of Starting Cohort 3 in Fifth Grade* [NEPS Working Paper No. 15]. Otto-Friedrich-Universität, Nationales Bildungspanel.

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

Rentzsch, K., Schröder–Abé, M., & Schütz, A. (2013). Being Called A 'Streber': The Roles of Personality and Competition in the Labelling of Academically Oriented Students. *European Journal of Personality*, *27*(5), 411–423. https://doi.org/10.1002/per.1884

Rentzsch, K., Schütz, A., & Schröder-Abé, M. (2011). Being Labeled*Nerd*: Factors That Influence the Social Acceptance of High-Achieving Students. *The Journal of Experimental Education*, *79*(2), 143–168. https://doi.org/10.1080/00220970903292900

Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The Power of Personality: The Comparative Validity of Personality Traits, Socioeconomic Status, and Cognitive Ability for Predicting Important Life Outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Robins, R. W., & Beer, J. S. (2001). Positive illusions about the self: Short-term benefits and long-term costs. *Journal of Personality and Social Psychology*, *80*(2), 340–352. https://doi.org/10.1037/0022-3514.80.2.340

Rosenberg, M. (2016). *Society and the Adolescent Self-Image*. Princeton University Press. https://doi.org/10.1515/9781400876136

Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion (2., vollst. überarb. und erw. Aufl.)*. Psychologie Lehrbuch. Huber.

Schimmack, U., & Kim, H. (2020). An integrated model of social psychological and personality psychological perspectives on personality and wellbeing. *Journal of Research in Personality*, *84*, 103888. https://doi.org/10.1016/j.jrp.2019.103888

Schnapp, T. (2020). *Samples, Weights, and Nonresponse: The Sample of Starting Cohort 3 of the National Educational Panel Study*. Leibniz Institute for Educational Trajectories, National Educational Panel Study.

Schnittjer, I., & Duchhardt, C. (2015). *Mathematical Competence: Framework and Exemplary Test Items*. Leibniz Institute for Science and Mathematics Education (IPN).

Schnittjer, I., & Gerken, A.-L. (2017). *NEPS Technical Report for Mathematics: Scaling Results of Starting Cohort 3 in Grade 7* [NEPS Survey Paper No. 16]. Leibniz Institute for Educational Trajectories, National Educational Panel Study.

Schönbrodt, F. D., & Humberg, S. (2021). *RSA: An R package for response surface analysis (Version 0.10.4)* [Computer software]. https://cran.r-project.org/package=RSA

Schütz, A., & Baumeister, R. F. (2017). Positive Illusions and the Happy Mind. In M. D. Robinson & M. Eid (Eds.), *The Happy Mind: Cognitive Contributions to Well-Being* (pp. 177–193). Springer International Publishing. https://doi.org/10.1007/978-3-319-58763-9_10

Schwartz, S. J., Coatsworth, J. D., Pantin, H., Prado, G., Sharp, E. H., & Szapocznik, J. (2006). The role of ecodevelopmental context and self-concept in depressive and externalizing symptoms in Hispanic adolescents. *International Journal of Behavioral Development*, *30*(4), 359–370. https://doi.org/10.1177/0165025406066779

Sedikides, C., & Gregg, A. P. (2008). Self-Enhancement: Food for Thought. *Perspectives on Psychological Science*, *3*(2), 102–116. https://doi.org/10.1111/j.1745-6916.2008.00068.x

Siddique, C. M., & D'Arcy, C. (1984). Adolescence, stress, and psychological well-being. *Journal of Youth and Adolescence*, *13*(6), 459–473. https://doi.org/10.1007/bf02088593

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford Univ. Press.

Snijders, T. A. B., & Bosker, R. J. (2011). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. SAGE.

Solon, G., Haider, S. J., & Wooldridge, J. M. (2015). What Are We Weighting For? *Journal of Human Resources*, *50*(2), 301–316. https://doi.org/10.3368/jhr.50.2.301

Stankov, L., & Crawford, J. D. (1997). Self-confidence and performance on tests of cognitive abilities. *Intelligence*, *25*(2), 93–109. https://doi.org/10.1016/s0160-2896(97)90047-7

Steinhauer, H. W., & Zinn, S. (2016). *NEPS Technical Report for Weighting: Weighting the Sample of Starting Cohort 3 of the National Educational Panel Study (Waves 1 to 5)*.

Steinmayr, R., Wirthwein, L., & Schöne, C. (2014). Gender and numerical intelligence: Does motivation matter? *Learning and Individual Differences*, *32*, 140–147. https://doi.org/10.1016/j.lindif.2014.01.001

Sugiura, N. (1978). Further analysts of the data by akaike' s information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, *7*(1), 13–26. https://doi.org/10.1080/03610927808827599

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*(2), 193–210. https://doi.org/10.1037/0033-2909.103.2.193

Vannatta, K., Gartstein, M. A., Zeller, M., & Noll, R. B. (2009). Peer acceptance and social behavior during childhood and adolescence: How important are appearance, athleticism, and academic competence? *International Journal of Behavioral Development*, *33*(4), 303–311. https://doi.org/10.1177/0165025408101275

Verschueren, K., Buyck, P., & Marcoen, A. (2001). Self-representations and socioemotional competence in young children: A 3-year longitudinal study. *Developmental Psychology*, *37*(1), 126–134. https://doi.org/10.1037/0012-1649.37.1.126

von Collani, G., & Herzberg, P. Y. (2003). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg. *Zeitschrift Für Differentielle Und Diagnostische Psychologie*, *24*(1), 3–7. https://doi.org/10.1024//0170-1789.24.1.3

Wagenmakers, E.-J., & Farrell, S. (2004). Aic model selection using Akaike weights. *Psychonomic Bulletin & Review*, *11*(1), 192–196. https://doi.org/10.3758/bf03206482

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

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54*(3), 427–450. https://doi.org/10.1007/bf02294627

Weinert, S., Artelt, C., Prenzel, M., Senkbeil, M., Ehmke, T., & Carstensen, C. H. (2011). 5 Development of competencies across the life span. *Zeitschrift Für Erziehungswissenschaft*, *14*(S2), 67–86. https://doi.org/10.1007/s11618-011-0182-7

Weinert, S. & NEPS, National Educational Panel Study. (2019). *NEPS-Startkohorte 3: Klasse 5 (SC3 8.0.0)* (8.0.0) [Data set]. NEPS National Education Panel Study. https://doi.org/10.5157/NEPS:SC3:8.0.0

Wohlkinger, F., Bayer, M., & Ditton, H. (2016). Measuring Self-Concept in the NEPS. In H.-P. Blossfeld, J. von Maurice, M. Bayer, & J. Skopek (Eds.), *Methodological issues of longitudinal surveys: The example of the National Educational Panel Study* (pp. 181–193). Springer VS. https://doi.org/10.1007/978-3-658-11994-2_11

Supplementary Materials

Peer Review History

Download: https://collabra.scholasticahq.com/article/37154-how-self-concept-competence-and-their-fit-or-misfit-predict-educational-achievement-well-being-and-social-relationships-in-the-school-context/attachment/94610.docx?auth_token=39Ow-bKcONnsc8KinA8d