Humanwissenschaftliche Fakultät

Jannis Bosch | Jürgen Wilbert

The impact of social comparison processes on self-evaluation of performance, self-concept, and task interest

Journal article | Version of record

The impact of social comparison processes on self-evaluation of performance, self-concept, and task interest

Jannis Bosch* and Jürgen Wilbert

Inclusive Education, University of Potsdam, Potsdam, Germany

Development of self-concept and task interest has been shown to be affected by social comparison processes in a variety of cross-sectional studies. A potential explanation for these effects is an effect of social comparative performance feedback on an individual's self-evaluation of performance, which in turn influences development of self-concept and task interest. There are, however, only few studies addressing this topic with experimental designs. This study was aimed at closing this research gap by experimentally manipulating social comparative performance. Feedback given was based on 2×2 experimental conditions: social position (high vs. low) and average performance of the reference group (high vs. low). Results show a strong effect of social position on self-evaluation of performance and smaller effects on self-concept and task interest.

KEYWORDS

social comparison, self-evaluation, self-concept, interest, contrast effects

Theory

Prior knowledge and general cognitive abilities are essential for the academic learning process. There are, however, several other motivational and affective characteristics of the learner that also shape the learning process in a substantial manner (Pintrich et al., 1993). Two constructs that have continuously been shown to be important for the learning process are self-concept and interest.

The self-concept can be conceptualized as a person's perceptions about themself in a particular domain (Marsh and Shavelson, 1985). In the academic domain, the most general dimension of self-concept is referred to as the *academic self-concept* (*ASC*), with more specific instances such as mathematics or reading self-concept further down the hierarchy. The development of the ASC is closely linked to academic achievement. It has been found to (a) be influenced by earlier academic achievement as well as (b) itself influencing future academic achievement (Guay et al., 2003; Marsh and Martin, 2011). While (a) can be explained by an increased success in academic tasks, leading to more positive external performance feedback as well as more positive self-evaluations, (b) can be explained by an increased ASC leading to enhanced confidence, effort and the use of deeper learning strategies (McInerney et al., 2012; Marsh et al., 2016).

Academic self-concept is also closely linked to academic interest. Interest closely resembles object-specific intrinsic motivation (Ryan and Deci, 2000) and is defined as the motivational orientation toward a specific object or domain without direct external rewards (Schiefele, 1992). Even though intrinsic motivation and interest are somewhat distinct in terms of research traditions, there is a strong overlap in the constructs. Within this manuscript, we will stick to the term interest. However, we will also include research that focusses on academic intrinsic motivation when the operationalization fits the presented definition of interest.

Associations between ASC and academic interest were consistently shown across a variety of studies and domains (e.g., van Kraayenoord and Schneider, 1999; Denissen et al., 2007; Korhonen et al., 2016; Lohbeck et al., 2016). The relation between self-concept and interest appears to be bidirectional, which can be explained by common sources affecting both self-concept and interest development. Denissen et al. (2007) suggest that the development of self-concept and interest within a domain is positively influenced by the understanding that one is capable of coping with problems in that domain. Theoretically, development of self-concept is thought to be shaped by the interpretation of one's experiences as well as the evaluations of important others such as friends or teachers (Shavelson et al., 1976). Based on previous research on intrinsic motivation, interest development can be linked to satisfaction of the basic psychological needs for autonomy, relatedness, and competence (Ryan and Deci, 2000; Krapp, 2005). According to Krapp (14), satisfaction of the need for competence in particular is based on the ability to attain valued outcomes when working on an interest-related task. A positive self-evaluation of academic performance is therefore necessary to maintain both a positive ASC and academic interest. A student's self-evaluation, in turn, is influenced by external evaluation structures (Ames, 1992). Therefore, it is worth investigating how exactly academic evaluation structures affect self-evaluation processes, as well as how these changes in self-evaluation affect the development of both self-concept and interest.

Most academic performance evaluation is based on social comparison processes. This focus on social comparisons subsequently affects students' self-evaluations and leads them to also focus on social comparison information (Dijkstra et al., 2008). The causal effect of social comparison information, such as grades, on both perceived competence and interest has been stressed by Cognitive Evaluation Theory (Deci and Ryan, 1985; Ryan and Deci, 2000) and could also be shown in several experimental studies. While Tauer and Harackiewicz (2004) found that a focus on interpersonal and intergroup competition can lead to an increase in both interest and performance, this effect seems to be influenced by the outcome of the competition. Reeve and Deci (1996), for example, reported that both interest and perceived competence were higher in participants that received bogus feedback informing them they won against another participant in a puzzle task compared to those that received feedback informing them they lost. Zell et al. (2017) replicated those findings and further elaborated them by showing that global (i.e., all other participants) and local (i.e., the last five participants) comparisons independently influence interest and perceived competence and that the effect of local comparison processes is comparatively stronger.

A detailed framework on *how* the processing of social comparison information affects self-evaluation processes can be found in the *Inclusion/Exclusion Model* (*IEM*; Schwarz and Bless, 1992; Bless and Schwarz, 2010). The authors suggest that the performance evaluation of a target is based on the mental representation of that target as well as the mental representation of a standard of comparison (here the performance of the class). New information about the standard of comparison can either be included into the mental representation of the target and therefore result in an assimilation effect (e.g., good or bad performance of the class is also attributed to oneself) or it can be excluded from the mental representation of the target and therefore result in a contrast effect (e.g., one's own performance is compared to the rest of the class).

Contrast effects are supposed to be the basis of regularly observed negative associations between average classroom performance and individual students' ASC. This association was described as the *Big-Fish-Little-Pond Effect* (*BFLPE*; Marsh and Parker, 1984; Marsh, 1987) and has since then been replicated in various school subjects in a large array of studies (e.g., Huguet et al., 2009; Jonkmann et al., 2012; Nagengast and Marsh, 2012). The BFLPE could also be shown to persist through almost all grade levels from third grade (Trautwein et al., 2008) to the final year of school (Trautwein et al., 2009; Jonkmann et al., 2012). Students with learning disorders in regular and special schools were also affected by the BFLPE in a similar way (Szumski and Karwowski, 2015).

Assimilation effects have also been investigated as an influence on the ASC under the label *Basking-In-Reflected-Glory Effect* (*BIRGE*). Here, a strong performing comparison group increases a person's ASC. However, in most studies this effect is overshadowed by the much stronger BFLPE. In order to investigate the BIRGE, researchers explicitly assessed the perceived school status and included it into statistical models (Marsh et al., 2000; Trautwein et al., 2009). This revealed a positive association between perceived school status and individual students' ASC, even when controlling for individual and school-average academic performance. To the author's knowledge, no studies investigating the effects of perceived school or class standing on academic interest have been conducted to date.

Another way to investigate the BIRGE is to compare classes where students were grouped by their ability. While Preckel and Brüll (2010) found evidence for a BIRGE comparing a sample of fifth-graders from regular classes to gifted classes, other authors neither found evidence for a BIRGE on ASC in a sample of ninth-graders of different ability-tracked school types nor in a sample of ninth-graders in different ability streams within one school (Trautwein et al., 2006). Even though the exact conditions for the occurrence of contrast and assimilation effects underlying BFLPE and BIRGE are not fully understood (Dai and Rinn, 2008), there is, nevertheless, clear evidence showing that the composition of the reference group influences ASC development.

These effects of reference group composition do not seem to be restricted to the ASC, however. Two studies showed that contrast effects similar to the BFLPE also influence the development of academic interest, though effects seem to be smaller than those on the ASC (Köller et al., 2000; Trautwein et al., 2006; Schurtz et al., 2014). Trautwein et al. (2006) also looked for the influence of potential assimilation effects on academic interest. Even though they did not find an assimilation effect on academic interest, their study was unable to replicate the BIRGE on ASC either, rendering these results somewhat ambiguous. The studies conducted under the labels of BFLPE and BIRGE support the idea that contrast and assimilation effects shape the ASC and interest of individual students in social comparison situations. They do not, however, directly target the effects of social comparative performance feedback on self-evaluation of performance. To directly investigate whether social comparison processes influence self-evaluation processes, experimental studies manipulating social comparative performance feedback, while directly assessing self-evaluation processes, are needed.

Based on presented empirical results and theoretical considerations, presenting a student with information about his or her own performance in comparison to the performance of a reference group, on the one hand, can be expected to result in contrast effects of self-evaluation (which subsequently affect self-concept and interest development). Presenting a student with information about the performance of a reference group in comparison to other groups, on the other hand, can be expected to result in assimilation effects of self-evaluation. These assimilation effects can also be expected to carry over to self-concept and interest development.

Compared to the large body of cross-sectional studies on BFLPE and BIRGE, fewer experimental studies have been conducted to investigate the causal mechanisms behind these phenomena. Zell and Alicke (2010), for example, manipulated social and individual temporal comparison feedback in a bogus social sensitivity task and found that both social and temporal comparison affected participants' self-rated social sensitivity. In a series of studies, Zell and Alicke (2009) could further show that bogus intra-group comparisons in a verbal reasoning task consistently influenced self-evaluations, while inter-group correlations only had an influence on self-evaluations when no intra-group comparison was available. In a similar vein, Zell et al. (2017) showed that local comparisons (i.e., comparisons to the last couple of participants) have a stronger influence than more global comparisons (i.e., comparisons to the average student of their university).

However, the only study experimentally manipulating the social position (i.e., whether participants performed better or worse than most of their peers) while also assessing post-feedback self-evaluation of performance, self-concept and interest was conducted by Pohlmann and Möller (2006). Participating university students received feedback about either a high (i.e., a percentile of around 88) or low (i.e., a percentile of around 23) social position within their reference group in two different academic learning tasks (a word analogy and a figure analogy task). The effects of social position on self-evaluation of performance were significant and showed strong effect sizes ($dcontrast = 1.88$ for the word analogy task and $dcontrast = 1.92$ for the figure analogy task), with participants receiving feedback about a high social position evaluating their own performance more positively. Effects on the task-specific self-concept were also significant and in the same direction, but with smaller effect sizes ($dcontrast = 0.87$ for the word analogy task and $dcontrast = 0.51$ for the figure analogy task). Task interest was only affected by experimental manipulation of the social position in the word analogy task ($dcontrast = 0.41$), while no effect could be found in the figure analogy task ($dcontrast = 0.00$). These results can be interpreted as evidence for the influence of contrast effects on self-evaluation of performance, and, to a lesser degree, also on self-concept and interest. The investigation of assimilation effects, however, was not within the scope of their study.

Bosch and Wilbert (2017, 2020) conducted two studies investigating the influence of experimentally manipulated social comparative performance feedback on task interest and self-evaluation of performance (the latter was only assessed in the second study, however). Both studies were using a 2×2 pre-post design manipulating feedback on both individual social position as well as absolute criterial score of the reference group. The first study (Bosch and Wilbert, 2017) found an effect of both the individual social position ($dcontrast = 0.22$) and the absolute criterial score of the reference group ($dassimilation = 0.20$) on task interest in a sample of 122 university students. The second study (Bosch and Wilbert, 2020) was able to replicate these effects on task interest in a sample of 230 elementary school children ($dcontrast = 0.30$ and $dassimilation = 0.27$) while showing an even stronger effect of both social position and criterial score of the reference group on self-evaluation of performance ($dcontrast = 1.26$ and $dassimilation = 0.74$). They did not assess post-manipulation self-concept, however, and the operationalization used to investigate potential assimilation effects (i.e., criterial score of the reference group) was confounded with the individual criterial score, making it impossible to clearly disentangle effects of the criterial score and genuine assimilation effects.

Effect sizes by Pohlmann and Möller (2006) were originally reported as Eta^2 but were converted to *Cohen's d* for better comparability (Fritz et al., 2012). Effect sizes for task interest of Bosch and Wilbert (2017, 2020) were calculated using the procedure suggested by Morris (2008) as *dppc2* on raw data of the respective study.

Research questions and hypotheses

To this day, there is no study investigating the influence of both contrast and assimilation effects on self-evaluation of performance, while also assessing potential consequences for self-concept and interest development in an experimental design. Most studies conducted on the topic have focused on the more distant construct of self-concept rather than directly assessing single instances of self-evaluation of performance. Because the self-concept is suggested to be the sum of past instances of self-evaluation of performance (Shavelson et al., 1976), these studies basically rely on the aggregated effects of many self-evaluations to show the effects of social comparison processes. Hence, the first goal of this study is to shed further light on the effects of social comparative performance feedback on self-evaluation of performance. In a similar vein, the second goal of this study is to investigate potential mediation effects of self-evaluation of performance on self-concept and task interest, providing a more detailed picture of the mechanisms suggested to be behind BFLPE and BIRGE.

Hypothesis 1

In accordance with previous results social comparative performance feedback can be expected to influence self-evaluation of performance (Pohlmann and Möller, 2006; Bosch and Wilbert, 2020). Hence, we expect both the position within the reference group as well as the relative performance of the reference group compared to other groups to be positively associated with self-evaluation of performance (i.e., a high position within the reference group and a high position of the reference group lead to a more positive self-evaluation of performance).

Hypothesis 2

Based on previously described theoretical considerations about the generation of the self-concept (Shavelson et al., 1976), we expect single instances of self-evaluation of performance to be positively associated with self-concept development. Based on this assumption and the assumptions outlined in hypothesis 1, we expect task specific self-concept to be influenced by social comparative performance feedback in a way similar to self-evaluation of performance (i.e., a high social position and a higher relative performance of the reference group lead to a more positive self-concept development). Additionally, we also expect these effects of social comparative performance feedback on self-concept development to be mediated by self-evaluation processes.

Hypothesis 3

Based on theoretical considerations about the development of interest (Deci and Ryan, 1985; Krapp, 2005) outlined earlier in the manuscript and on the empirically observed associations between self-concept and interest (Denissen et al., 2007), we expect social comparative performance feedback to have a similar effect on task interest. Hence, similar to hypothesis 2, we also expect effects of

performance feedback on task interest to be mediated by self-evaluation processes.

Methods

Participants

Participants were 190 first-year University of Potsdam students pursuing a degree in elementary school education. Five participants were excluded because they stated University of Potsdam was not the University they primarily identified with, 19 participants were excluded because they expressed doubt about the authenticity of the performance feedback during the manipulation check, leaving us with a sample of 166 participants for further analyses. 147 participants were female, and 19 participants were male. Participants' age ranged from 18 to 46 ($M = 22.48$, $SD = 5.65$).

Measures

University identification

Four items were used to assess university identification (e.g., "I feel I have a lot in common with other students affiliated with University of Potsdam"). Items were based on the university identification scale used by Cho and Yu (2015) and translated into German. Reliability of the scale was acceptable (*Cronbach's* $\alpha = 0.72$).

Self-concept

Self-concept was assessed using four items: two positive (e.g., "I always did well on memory tasks") and two negative items (e.g., "Dealing with learning tasks is not one of my strengths"). The items were based on the math self-concept scale used by Schwanzer et al. (2005) and adapted to refer to the learning task used in this study. The self-concept scale was administered twice: once after finishing the instruction (pre-test; T1) and once after the second run of the learning task (post-test; T2). Participants rated each item on a Likert-scale ranging from 1 (strongly disagree) to 7 (strongly agree). Both pre- and post-test self-concept showed high internal consistency (*Cronbach's* $\alpha pre = 0.91$ and *Cronbach's* $\alpha post = 0.90$).

Task interest

To measure task interest, a scale consisting of four items (e.g., "I like learning tasks such as this one") was used. The task interest scale was based on an interest scale in a German motivation questionnaire (Rheinberg et al., 2001). Items were slightly adapted to refer to the learning task used in this study. Similar to the self-concept scale, the task interest scale was administered twice (T1 and T2) and items were rated on a Likert-scale ranging from 1 (strongly disagree) to 7 (strongly agree). Both pre- and post-test task interest showed high internal consistency (*Cronbach's* $\alpha pre = 0.92$ and *Cronbach's* $\alpha post = 0.94$).

Self-evaluation of performance

Self-evaluation of performance was measured directly after receiving experimentally manipulated feedback following the first run of the learning game. Participants were asked to rate their own performance during the previous run on a Likert-scale from 1 (very bad) to 5 (very good).
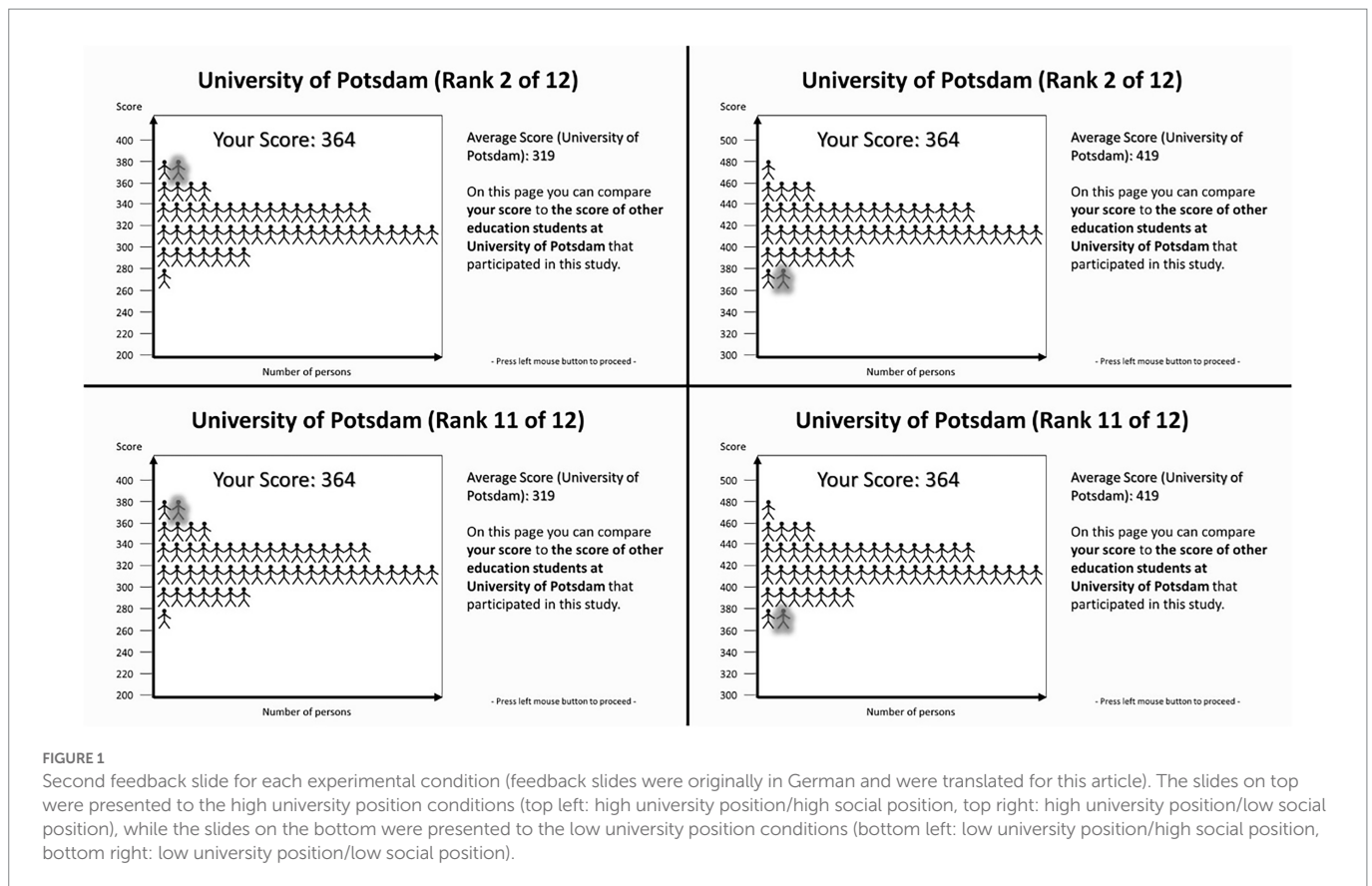
Learning task

The learning task used in this study was an adaptation of "*The Flag Game*," a learning game used in an earlier study on a similar topic (Bosch and Wilbert, 2017). The game consists of two separate phases: a learning and a performance phase. The learning phase consists of 30 trials, during each trial participants were presented with a flag and a map of the African continent, where the corresponding country outlines where highlighted. They were told to memorize as many combinations as possible. Each combination of country outlines and national flag was presented for a maximum of 7 s. Once they finished the 30 items from the learning phase, they moved on to the performance phase. During the performance phase, the same 30 national flags were presented to the participants, but this time with a selection of five different highlighted country outlines (i.e., the corresponding country outline and four distractors). Their task was to identify the correct combination as quickly as possible. Participants were told their score was based on both accuracy and speed of their answers in the performance phase to make it harder for them to judge their own performance. After the first run (i.e., a learning and a performance phase) participants received experimentally manipulated feedback about their task performance. Then a second run with the same 30 items followed, but this time no feedback was given.

To make sure there were no sequencing effects, the trial order was randomized separately for each learning and performance phase. Distractors items were taken from the original item pool, balanced and pseudo-randomized separately for both performance phases. Hence, each participant was presented with the same sequence of items and distractors.

Feedback conditions

During the instruction, participants were told they would receive information about both their own performance in comparison to other students from their own university (social position) as well as information about the average performance of students of their own university compared to students of other universities (university position). Immediately after the first run of the flag game, participants were presented with a feedback slide containing information about the average performance of students of their own university compared to students of other universities. There were two different university position conditions: either students of their university performed very well (i.e., second out of 12 universities) or very badly (i.e., eleventh out of 12 universities). Afterward, they received a second feedback slide with information about their own performance compared to fellow students at their university (see Figure 1). Similar to the university position conditions, there were two different social position conditions: performance feedback either showed participants placing in the top or the bottom 5% of their peers. Hence, to test for potential contrast or assimilation effects performance feedback was given based on 2 (high and low university position; *UP+/UP-*) × 2 (high and low social position; *SP+/SP-*) experimental conditions. There were 38 participants in the SP-/UP- condition, 38 participants in the SP+/UP- condition, 44 participants in the SP-/UP+ condition, and 46 in the SP+/UP+ condition.

To control for potential effects of the criterial score, each participant, independent of experimental condition, received a score of 364 points.

FIGURE 1
Second feedback slide for each experimental condition (feedback slides were originally in German and were translated for this article). The slides on top were presented to the high university position conditions (top left: high university position/high social position, top right: high university position/low social position), while the slides on the bottom were presented to the low university position conditions (bottom left: low university position/high social position, bottom right: low university position/low social position).

Hence, the average score of other students of University of Potsdam was varied depending on the social position condition: participants in the high social position conditions were told the average score of their peers was 319 points while participants in the low social position conditions were told the average score of their peers was 419 points (see Figure 1).

Procedure

All participants were recruited during a first-year inclusive education lecture. They were given a link to an online platform where they were able to sign up for any available appointment. Participants were tested in groups of 6–10 persons. Prior to testing, written informed consent was obtained from each participant. After a short introduction by the test supervisor, each participant was placed in front of a visually shielded computer. The test started with a short computer-based instruction about the learning game and the upcoming procedure, participants were then asked to fill out the university identification and T1 self-concept and task interest questionnaires, as well as a questionnaire about social comparison orientation (Gibbons and Buunk, 1999) not used within the scope of this manuscript. Participants then proceeded with the first run of the learning game followed by the experimentally manipulated performance feedback. They were then asked to evaluate their own performance during the first run. Afterward, the second run of the learning game was started followed by T2 self-concept and task interest questionnaires. Participants then answered two questions concerning their interest in similar future studies and an open question that was used as a manipulation check asking whether participants noticed anything unusual during the study. The open question was separately evaluated by two researchers to identify students

who had doubts about the authenticity of the feedback intervention. Students who were classified as suspicious by both researchers were subsequently excluded from further analyses (see participants section). Finally, the true nature of the experiment was uncovered, and the experimental manipulation was explained to all participants.

Statistical analyses

All statistical calculations were carried out using the statistical software R (R Core Team, 2019).

To test hypothesis 1, suggesting a positive influence of high social and university position on self-evaluation of performance, two linear regression models were calculated, with self-evaluation of performance as criterion. Model 1 only contained the intercept. For model 2a and 2b, social position and university position were included, respectively. Finally, model 3 included both social position and university position as well as their interaction. Additionally, self-evaluation of performance was compared between participants in the incongruent conditions (i.e., low social position/high university position and high social position/low university position).

To test hypotheses 2 and 3, linear mixed models with fixed and random effects were used in order to be able to accommodate the repeated measurements of both variables. Hence, self-concept and interest were used as criterion in separate analyses, but similar predictors were added for each criterion. Model 1 only contained the intercept and measurement time as predictors. Model 2a included social position and its interaction with measurement time while model 2b included university position and its interaction with measurement time. In model 3 both experimental conditions (i.e., social position, university position

and their interaction) and their interaction with measurement time were included (*lme4* package for R; Bates et al., 2015). Additionally, self-concept and interest development were compared between participants in the incongruent conditions.

Further, causal mediation analyses were carried out with task self-concept and task interest (post-test) as respective outcome variables, self-evaluation of performance as mediator, experimental conditions as treatment variable, and self-concept and task interest (pre-test) as respective control variables. Because participants were randomly assigned to respective experimental conditions, model-based causal mediation analyses were conducted (*mediation* package for R; Tingley et al., 2014), as suggested by Imai et al. (2010).

Results

Table 1 shows the intercorrelations between all relevant study variables. As expected, T1 self-concept shows a moderate correlation to T1 task interest. Further, T1 task interest and task performance also show a small, positive correlation. Hence, expected positive associations between initial self-concept and task interest could be replicated in this sample. The same pattern could be found for the respective T2 variables.

Self-evaluation of performance shows no significant correlation to T1 self-concept, but a positive correlation to T2 self-concept. This is not surprising, since both T2 self-concept and self-evaluation of performance were assessed after the manipulated performance feedback, while T1 self-concept was not. Table 2 shows the descriptive results separately for each group as well as for the full sample.

Hypothesis 1

Effects of experimental conditions on self-evaluation of performance

Table 3 shows the linear regression models predicting self-evaluation of performance. While model 2a does increase variance explained compared to model 1 [$\Delta R^2 = 0.50$, $F(1, 164) = 168.40$, $p < 0.001$], model 2b does not [$\Delta R^2 = 0.00$, $F(1, 164) = 0.06$, $p = 0.816$]. Model 3, which contains both experimental conditions and their interaction as predictors, significantly increases variance explained compared to model 1 [$\Delta R^2 = 0.51$, $F(3, 162) = 58.29$, $p < 0.001$]. A closer look at the regression weights shows that social position as well as the interaction social position × university position significantly predict self-evaluation of performance, while university position itself does not. Differences between participants in the incongruent conditions were significant [$\Delta R^2 = 0.54$, $F(1, 80) = 93.79$, $p < 0.001$]. Results did not change substantially when controlling for actual performance, university identification or social comparison orientation.

To shed more light on the nature of the interaction, *post-hoc* one-factorial linear regressions were conducted to see whether there are differential results of social position in the respective high and low university position conditions and the other way around. Results show a significant effect of social position in both the low university position [$R^2 = 0.41$, $F(1, 74) = 52.52$, $p < 0.001$] and the high university position [$R^2 = 0.58$, $F(1, 88) = 123.50$, $p < 0.001$] conditions, while university position only showed a trend toward significance in the low social position [$R^2 = 0.03$, $F(1, 80) = 3.88$, $p = 0.052$] and no significant effect in the high social position [$R^2 = 0.01$, $F(1, 82) = 1.60$, $p = 0.21$] condition.

Hence, hypothesis 1 can only be partially confirmed. While a high social position clearly has positive effects on self-evaluation of performance, the results for university position are less clear cut.

Hypothesis 2

Effects of experimental conditions on development of self-concept

Table 4 (self-concept) shows model and parameter estimates for linear mixed models predicting development of self-concept. Model 1 shows the base model, only containing measurement time as predictor. Model 2a is the social position model, containing social position, measurement time, and their interaction as predictors. Model 2b is the university position model, containing university position, measurement time and their interaction as predictors. In Model 3, social position, university position, and the interaction social position × university position as well as all possible interactions with measurement time were added to the base model. Both model 2a ($\Delta R^2 Beta = 0.056$, *Likelihood Ratio* = 21.31, $p < 0.001$) and model 3 ($\Delta R^2 Beta = 0.092$, *Likelihood Ratio* = 27.23, $p < 0.001$) significantly increased model fit compared to model 1, while model 2b ($\Delta R^2 Beta = 0.022$, *Likelihood Ratio* = 2.37, $p > 0.05$) did not. Results did not change substantially when controlling for actual performance, university identification or social comparison orientation. Further, differences in self-concept development between the incongruent conditions were also significant ($\Delta R^2 Beta = 0.016$, *Likelihood Ratio* = 7.35, $p < 0.05$).

Measurement time emerged as a significant negative predictor in all models, suggesting there was a negative development of self-concept from pre- to post-test. This negative development, however, was significantly influenced by the positive interaction measurement time × social position in models 2a and 3. Hence, participants in the high social position conditions did not show the negative development in task self-concept observed in the low social position conditions (see Table 2). However, contrary to our hypothesis, there was no positive interaction measurement time x university position, suggesting university position as presented in this study did not influence self-concept development.

Causal mediation analysis for self-concept development

As can be seen in Table 1, self-evaluation of performance shows a positive correlation with task-specific self-concept development from pre- to post-test ($r = 0.34$, $p < 0.01$). Results of the mediation analysis further show a significant *average causal mediation effect* ($ACME = 0.21$, $p < 0.01$) with 60% of the effect on self-concept development explained by self-evaluation of performance (*proportion mediated; PM* = 0.60, $p < 0.01$), a non-significant *average direct effect* ($ADE = 0.14$, $p = 0.21$), and a significant *total effect* ($TE = 0.34$, $p < 0.001$). Hence, while there is a significant causal mediation by self-evaluation of performance, the direct effect of the social position condition does not have a significant influence on self-concept development.

Hypothesis 3

Effects of experimental conditions on development of task interest

Table 4 (task interest) shows model and parameter estimates for linear mixed models predicting development of task interest used to test

TABLE 1 Intercorrelations for all observed variables (*N*=166).

	1.	2.	3.	4.	5.	6.	7.
1. SC T1	1						
2. SC T2	0.86***	1					
3. SC DIF	−0.22**	0.31***	1				
4. INT T1	0.42***	0.46***	0.09	1			
5. INT T2	0.41***	0.47***	0.14	0.90***	1		
6. INT DIF	0.05	0.12	0.13†	−0.04	0.41**	1	
7. SEoP	0.12	0.30***	0.34***	0.07	0.13†	0.16*	1
8. ID	0.03	0.00	−0.06	0.20*	0.18*	0.00	−0.05

***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, †$p < 0.10$; SC, self-concept; INT, interest; SEoP, self-evaluation of performance; ID, identification; T1, pre-intervention; T2, post-intervention; and DIF, difference between T1 and T2.

TABLE 2 Descriptive statistics for all observed variables separately for each experimental condition (*N*=166).

Variable	SP+/UP+	SP+/UP-	SP-/UP+	SP-/UP-	Total
MT	Mean (SD)				
Identification					
only T1	4.54 (0.90)	4.53 (1.03)	4.68 (1.08)	4.82 (0.97)	4.64 (1.00)
Self-Concept					
T1	4.52 (1.18)	3.99 (1.26)	4.10 (1.15)	4.14 (1.39)	4.20 (1.25)
T2	4.50 (1.18)	4.07 (1.27)	3.80 (1.16)	3.64 (1.38)	4.02 (1.28)
Difference	−0.02 (0.55)	0.08 (0.58)	−0.31 (0.69)	−0.51 (0.73)	−0.19 (0.67)
Interest					
T1	4.39 (1.38)	4.00 (1.42)	4.53 (1.43)	4.20 (1.59)	4.29 (1.46)
T2	4.53 (1.51)	4.07 (1.65)	4.40 (1.60)	3.99 (1.62)	4.26 (1.59)
Difference	0.15 (0.68)	0.07 (0.74)	−0.12 (0.58)	−0.22 (0.80)	−0.03 (0.71)
Self-evaluation of performance					
only T2	3.54 (1.13)	3.24 (1.08)	1.48 (0.51)	1.74 (0.69)	2.51 (1.27)

SP, social position; UP, university position; MT, measurement time; T1, pre-intervention; and T2, post-intervention.

TABLE 3 Fixed effects for the linear regression model predicting self-evaluation of performance (*N*=166).

Parameter	Stand. Estimate Beta (SE)			
	Model 1	Model 2a	Model 2b	Model 3
Intercept	2.51*** (0.10)	1.60*** (0.10)	2.49*** (0.15)	1.74*** (0.15)
SP		1.81*** (0.14)		1.50*** (0.20)
UP			0.05 (0.20)	−0.26 (0.20)
SP × UP				0.57* (0.28)
Model indices				
df	165	164	164	162
Adj. R²	0.00	0.51	0.00	0.51

***$p < 0.001$, *$p < 0.05$, †$p < 0.10$; SP, social position (0/1); UP, university position (0/1).

hypothesis 3. Once again, model 1 is the base model only containing measurement time. Model 2a is the social position model, containing social position, measurement time as well as their interaction as

predictors. Model 2b contains the university position model with university position, measurement time and their interaction as predictors. Model 3 once again contained measurement time, social position, and university position, as well as all their interactions. Only model 2a significantly increased variance explained ($\Delta R^2 Beta = 0.005$, *Likelihood Ratio* = 6.75, $p < 0.05$) compared to model 1, while model 2b ($\Delta R^2 Beta = 0.038$, *Likelihood Ratio* = 3.61, $p = 0.17$) and model 3 ($\Delta R^2 Beta = 0.043$, *Likelihood Ratio* = 10.37, $p = 0.11$) did not. Results did not change substantially when controlling for actual performance, university identification or social comparison orientation.

In model 1, measurement time does not emerge as a significant predictor. This means there was no change in task interest from pre- to post-test across all experimental conditions. However, in model 2a both measurement time and the interaction measurement time × social position are significant predictors. This pattern of results suggests that while participants in the low social position conditions experienced a decrease of task interest, participants in the high social position conditions experienced an increase in task interest. Model 2b does not include any significant effects, suggesting that the university position condition did not affect task interest. In model 3, a pattern of results similar to model 2a can be found. However, this time both measurement time and the

TABLE 4 Fixed effects for mixed models predicting self-concept and task interest (*N* =166)

Parameter	Stand. Estimate Beta (SE)							
	Self-concept				Task interest			
	Model 1	Model 2a	Model 2b	Model 3	Model 1	Model 2a	Model 2b	Model 3
Intercept	4.20*** (0.10)	4.12*** (0.14)	4.07*** (0.15)	4.15*** (0.20)	4.29*** (0.12)	4.38*** (0.17)	4.10*** (0.17)	4.20*** (0.25)
Level 1								
MT	−0.19*** (0.05)	−0.40*** (0.07)	−0.21** (0.08)	−0.51*** (0.10)	−0.03 (0.06)	−0.17* (0.06)	−0.07 (0.08)	−0.22† (0.11)
Level 2								
SP		0.16 (0.19)		−0.15 (0.29)		−0.17 (0.24)		−0.20 (0.35)
UP			0.25 (0.20)	−0.04 (0.28)			0.35 (0.24)	0.33 (0.34)
Interactions								
MT × SP		0.42*** (0.10)		0.59*** (0.15)		0.28** (0.11)		0.29† (0.16)
MT × UP			0.05 (0.11)	0.20 (0.14)			0.09 (0.11)	0.09 (0.15)
SP × UP				0.57 (0.39)				0.06 (0.48)
MT × SP × UP				−0.30 (0.20)				−0.02 (0.22)
Model indices								
AICc	882	865	884	867	964	960	964	965
df	4	6	6	10	4	6	6	10
R^2Beta	0.011	0.067	0.033	0.103	0.000	0.005	0.038	0.043

***$p < 0.001$, †$p < 0.10$; MT, measurement time (0/1); SP, social position (0/1); UP, university position (0/1); AICc, corrected Akaike Information Criterion; and df, degrees of freedom.

interaction measurement time x social position only show marginally significant effects. Differences between incongruent conditions were not statistically significant ($\Delta R^2Beta = 0.050$, *Likelihood Ratio* = 3.59, $p = 0.17$).

Causal mediation analysis for task interest development

Self-evaluation of performance shows a small but significant positive correlation with task-specific interest development from pre- to post-test ($r = 0.16$, $p < 0.05$). In order to investigate whether self-evaluation of performance mediates the effect of social position on interest development in a similar fashion to self-concept development, another causal mediation analysis was conducted. However, this time results show a non-significant *average causal mediation effect* ($ACME = 0.03$, $p = 0.70$) with a non-significant *proportion mediated* ($PM = 0.13$, $p = 0.70$), a non-significant *average direct effect* ($ADE = 0.15$, $p = 0.15$), and a significant *total effect* ($TE = 0.17$, $p < 0.05$). This shows that self-evaluation of performance does not mediate the effect of social position on task interest development.

Discussion

The first goal of this study therefore was to closely investigate contrast and assimilation effects of social comparative performance feedback on self-evaluation of performance. As hypothesized, results show that persons receiving feedback about a high social position evaluate themselves more positively than persons receiving feedback about a low social position, suggesting participants used the presented reference group to contrast against their own performance. The effect of experimentally manipulated social comparative performance feedback on self-evaluation of performance was relatively strong, explaining about half of the latter's variance. Hence,

these results provide evidence for the existence of a contrast effect. Contrary to our expectations, however, there was no evidence suggesting an assimilation effect on self-evaluation of performance. Hence, the mechanism suggested to be behind assimilation effects on self-evaluation of performance found in several observational studies could not be supported by the data of this particular study. The lack of an assimilation effect, however, might be due to several reasons: Firstly, it might be caused by a lack of identification of the students with University of Potsdam. According to Hall and Crisp (2008), ingroup identification is a necessary prerequisite for the occurrence of assimilation effects. Secondly, the performance in a learning task, such as the flag game, might not be able to influence the perceived status or reputation of the university sufficiently to invoke a measurable assimilation effect. Thirdly, the learning task presented to study participants might not be sufficiently connected to actual content learned at university. Finally, the absence of assimilation effects might also be caused by a mechanism referred to as contextual neglect (Zell and Alicke, 2009), i.e., participants tend to ignore intergroup level feedback in their self-evaluation of performance when intragroup feedback is available. Hence, additional experimental research is necessary to clear up whether the BIRGE found in several cross-sectional studies (e.g., Marsh et al., 2000; Trautwein et al., 2009) is based on assimilation effects as suggested by the authors. Further research on situational factors contributing to the presence or absence of assimilation effects is also necessary to acquire a deeper understanding of the processes involved.

The second goal of this study was to investigate subsequent effects of social comparative performance feedback on task interest and self-concept. Results of the present study suggest that the social position presented to study participants during social comparative performance feedback influenced the development of self-concept and task interest

in a similar manner to self-evaluation of performance, but with smaller effect size. The effect sizes on self-evaluation of performance on the one hand and task interest and ASC on the other, are not directly comparable, however. Because self-evaluation of performance is based on the external evaluation (i.e., performance feedback), it was assessed immediately after receiving feedback. Task interest and ASC, however, are not directly based on a single instance of performance feedback, but on the general attitudes and cognitions toward the respective domain. That is why they were assessed after another run of the flag game. Hence, the relatively longer time between feedback and assessment could have also caused the decrease in effect sizes.

Additionally, this study was designed to test whether the effects of experimental conditions on task interest and self-concept are mediated by participants' self-evaluation of performance. As expected in several theories (e.g., Shavelson et al., 1976; Deci and Ryan, 1985; Krapp, 2002), self-evaluation of performance was positively associated with the development of participants' self-concept as well as task interest over the duration of the study, which can be interpreted as evidence that a higher self-evaluation of performance leads to a more positive development of both self-concept and task interest. Accordingly, results of causal mediation analyses suggest that effects of the social position condition on self-concept development were substantially mediated by self-evaluation of performance. Effects of the social position condition on task interest development, however, were not mediated by self-evaluation of performance. Therefore, results of this study provide clear evidence that single instances of self-evaluation are directly linked to self-concept development and that both self-evaluation and self-concept development are influenced by social comparison processes. Mechanisms behind the influences of social comparison processes on task interest development are not as clear, however. While we could find a small positive association between self-evaluation of performance and task interest development, results of the causal mediation analysis do not support the hypothesis that effects of social position on task interest are mediated by self-evaluation processes.

In summary, the results of this study provide experimental evidence for the hypothesized mechanisms behind the BFLPE, suggesting that the BFLPE found in a large array of studies (e.g., Marsh and Hau, 2003; Trautwein et al., 2006) is caused by contrast-effects of self-evaluation, which in turn affects self-concept development. Task interest appears to be less strongly associated with self-evaluation processes but is still affected by social comparison processes. The hypothesized assimilation effects on self-evaluation of performance could not be found, however.

Practical implications

As presented earlier in the manuscript, both ASC and task interest play a vital role in the academic learning process (e.g., Hidi, 1990; Korhonen et al., 2016; Oberle, 2018). Results of this study suggest that social comparative performance feedback clearly benefits already high performing students, while those placing at the bottom of the reference group show a less favorable development, at least in terms of ASC and task interest. This might not necessarily be problematic in a classroom where the relative academic performance between students is relatively homogeneous. However, it can be expected to lead to growing affective and motivational differences between the best and worst performing students in classrooms with heterogeneous performance levels where students with a comparatively low academic performance are not able to catch up with their better performing classmates. Hence, the effects of our current evaluation systems on the affective, motivational and cognitive development of every student needs to be put under scrutiny, particularly when the current shift in the education system is set out to increase heterogeneity in the average classroom.

Limitations

The very controlled and experimental nature of this study did provide several advantages when it comes to internal validity. However, there are several aspects of regular classroom interactions that had to be neglected in order to create the controlled environment used in this study. Hence, while the results of this individual study are to be interpreted with caution, particularly when directly inferring adaptations on the classroom level, the combined results of observational studies with relatively high external validity (e.g., studies under the label of the BFLPE) and experimental studies like the one presented in this paper still provide a differentiated picture of the effects of social comparison processes that can be used to adapt educational evaluation procedures, particularly those with a strong focus on social comparison.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/V4WGX/.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

Author contributions

The study was planned and conceptualized by JB and JW. Project administration, software programming, and original draft preparation were done by JB. JW provided supervision during the whole project. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ames, C. (1992). Classrooms: goals, structures, and student motivation. *J. Educ. Psychol.* 84, 261–271. doi: 10.1037/0022-0663.84.3.261

Bates, D., Mächler, M., and Bolker, B. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bless, H., and Schwarz, N. (2010). Mental construal and the emergence of assimilation and contrast effects. *Adv. Exp. Soc. Psychol.* 42, 319–373. doi: 10.1016/S0065-2601(10)42006-7

Bosch, J., and Wilbert, J. (2017). Contrast and assimilation effects on task interest in an academic learning task. *Frontline Learn. Res.* 5, 60–77. doi: 10.14786/flr.v5i2.292

Bosch, J., and Wilbert, J. (2020). Contrast and assimilation effects on self-evaluation of performance and task interest in a sample of elementary school children. *Front. Educ.* 4:165. doi: 10.3389/feduc.2019.00165

Cho, J., and Yu, H. (2015). Roles of university support for international students in the United States: analysis of a systematic model of university identification, university support, and psychological well-being. *J. Stud. Int. Educ.* 19, 11–27. doi: 10.1177/1028315314533606

Dai, D. Y., and Rinn, A. N. (2008). The big-fish-little-pond effect: what do we know and where do we go from here? *Educ. Psychol. Rev.* 20, 283–317. doi: 10.1007/s10648-008-9071-x

Deci, E. L., and Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.

Denissen, J. J. A., Zarrett, N. R., and Eccles, J. S. (2007). I like to do it, I'm able, and I know I am: longitudinal couplings between domain-specific achievement, self-concept, and interest. *Child Dev.* 78, 430–447. doi: 10.1111/j.1467-8624.2007.01007.x

Dijkstra, P., Kuyper, H., van der Werf, G., Buunk, A. P., and van der Zee, Y. G. (2008). Social comparison in the classroom: a review. *Rev. Educ. Res.* 78, 828–879. doi: 10.3102/0034654308321210

Fritz, C. O., Morris, P. E., and Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *J. Exp. Psychol. Gen.* 141, 2–18. doi: 10.1037/a0024338

Gibbons, F. X., and Buunk, B. P. (1999). Individual differences in social comparison: development of a scale of social comparison orientation. *J. Pers. Soc. Psychol.* 76, 129–142. doi: 10.1037/0022-3514.76.1.129

Guay, F., Marsh, H. W., and Boivin, M. (2003). Academic self-concept and academic achievement: developmental perspectives on their causal ordering. *J. Educ. Psychol.* 95, 124–136. doi: 10.1037/0022-0663.95.1.124

Hall, N. R., and Crisp, R. J. (2008). Assimilation and contrast to group primes: the moderating role of ingroup identification. *J. Exp. Soc. Psychol.* 44, 344–353. doi: 10.1016/j.jesp.2007.07.007

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Rev. Educ. Res.* 60:549. doi: 10.2307/1170506

Huguet, P., Dumas, F., Marsh, H., Régner, I., Wheeler, L., Suls, J., et al. (2009). Clarifying the role of social comparison in the big-fish–little-pond effect (BFLPE): an integrative study. *J. Pers. Soc. Psychol.* 97, 156–170. doi: 10.1037/a0015558

Imai, K., Keele, D., Tingley, D., and Yamamoto, T. (2010). "Causal mediation analysis using R," in *Advances in Social Science Research Using R.* ed. H. D. Vinod (New York: Springer), 129–154.

Jonkmann, K., Becker, M., Marsh, H. W., Lüdtke, O., and Trautwein, U. (2012). Personality traits moderate the big-fish–little-pond effect of academic self-concept. *Learn. Individ. Differ.* 22, 736–746. doi: 10.1016/j.lindif.2012.07.020

Köller, O., Schnabel, K. U., and Baumert, J. (2000). Der Einfluß der Leistungsstärke von Schulen auf das fachspezifische Selbstkonzept der Begabung und das Interesse. *Z. Entwicklungspsychol. Padagog Psychol.* 32, 70–80. doi: 10.1026//0049-8637.32.2.70

Korhonen, J., Tapola, A., Linnanmäki, K., and Aunio, P. (2016). Gendered pathways to educational aspirations: the role of academic self-concept, school burnout, achievement and interest in mathematics and reading. *Learn. Instr.* 46, 21–33. doi: 10.1016/j.learninstruc.2016.08.006

Krapp, A. (2002). Structural and dynamic aspects of interest development: theoretical considerations from an ontogenetic perspective. *Learn. Instr.* 12, 383–409. doi: 10.1016/S0959-4752(01)00011-1

Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learn. Instr.* 15, 381–395. doi: 10.1016/j.learninstruc.2005.07.007

Lohbeck, A., Nitkowski, D., and Petermann, F. (2016). A control-value theory approach: relationships between academic self-concept, interest, and test anxiety in elementary school children. *Child Youth Care Forum* 45, 887–904. doi: 10.1007/s10566-016-9362-1

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *J. Educ. Psychol.* 79, 280–295. doi: 10.1037/0022-0663.79.3.280

Marsh, H. W., and Hau, K.-T. (2003). Big-fish-little-pond effect on academic self-concept. A cross-cultural (26-country) test of the negative effects of academically selective schools. *Am. Psychol.* 58, 364–376. doi: 10.1037/0003-066X.58.5.364

Marsh, H. W., Kong, C.-K., and Hau, K.-T. (2000). Longitudinal multilevel models of the big-fish-little-pond effect on academic self-concept: counterbalancing contrast and reflected-glory effects in Hong Kong schools. *J. Pers. Soc. Psychol.* 78, 337–349. doi: 10.1037//0022-3514.78.2.337

Marsh, H. W., and Martin, A. J. (2011). Academic self-concept and academic achievement: relations and causal ordering: academic self-concept. *Br. J. Educ. Psychol.* 81, 59–77. doi: 10.1348/000709910X503501

Marsh, H. W., and Parker, J. W. (1984). Determinants of student self-concept: is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *J. Pers. Soc. Psychol.* 47, 213–231. doi: 10.1037/0022-3514.47.1.213

Marsh, H. W., Pekrun, R., Lichtenfeld, S., Guo, J., Arens, A. K., and Murayama, K. (2016). Breaking the double-edged sword of effort/trying hard: developmental equilibrium and longitudinal relations among effort, achievement, and academic self-concept. *Dev. Psychol.* 52, 1273–1290. doi: 10.1037/dev0000146

Marsh, H. W., and Shavelson, R. (1985). Self-concept: its multifaceted, hierarchical structure. *Educ. Psychol.* 20, 107–123. doi: 10.1207/s15326985ep2003_1

McInerney, D. M., Cheng, R. W., Mok, M. M. C., and Lam, A. K. H. (2012). Academic self-concept and learning strategies: direction of effect on student academic achievement. *J. Adv. Acad.* 23, 249–269. doi: 10.1177/1932202X12451020

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organ. Res. Methods* 11, 364–386. doi: 10.1177/1094428106291059

Nagengast, B., and Marsh, H. W. (2012). Big fish in little ponds aspire more: mediation and cross-cultural generalizability of school-average ability effects on self-concept and career aspirations in science. *J. Educ. Psychol.* 104, 1033–1053. doi: 10.1037/a0027697

Oberle, E. (2018). Early adolescents' emotional well-being in the classroom: the role of personal and contextual assets. *J. Sch. Health* 88, 101–111. doi: 10.1111/josh.12585

Pintrich, P. R., Marx, R. W., and Boyle, R. A. (1993). Beyond cold conceptual change: the role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Rev. Educ. Res.* 63, 167–199. doi: 10.3102/00346543063002167

Pohlmann, B., and Möller, J. (2006). Vergleichseffekte auf kognitive, affektive und motivationale Variablen. *Z. Entwicklungspsychol. Padagog. Psychol.* 38, 79–87. doi: 10.1026/0049-8637.38.2.79

Preckel, F., and Brüll, M. (2010). The benefit of being a big fish in a big pond: contrast and assimilation effects on academic self-concept. *Learn. Individ. Differ.* 20, 522–531. doi: 10.1016/j.lindif.2009.12.007

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available at: https://www.R-project.org/

Reeve, J., and Deci, E. L. (1996). Elements of the competitive situation that affect intrinsic motivation. *Personal. Soc. Psychol. Bull.* 22, 24–33. doi: 10.1177/0146167296221003

Rheinberg, F., Vollmeyer, R., and Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica* 47, 57–66. doi: 10.1026//0012-1924.47.2.57

Ryan, R. M., and Deci, E. L. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemp. Educ. Psychol.* 25, 54–67. doi: 10.1006/ceps.1999.1020

Schiefele, U. (1992). "Topic interest and leveles of text comprehension," in *The Role of Interest in Learning and Development*. eds. K. A. Renninger, S. Hidi and A. Krapp (Hillsdale, New Jersey: L. Erlbaum Associates), 151–182.

Schurtz, I. M., Pfost, M., Nagengast, B., and Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learn. Instr.* 34, 32–41. doi: 10.1016/j.learninstruc.2014.08.001

Schwanzer, A. D., Trautwein, U., Lüdtke, O., and Sydow, H. (2005). Entwicklung eines Instruments zur Erfassung des Selbstkonzepts junger Erwachsener. *Diagnostica* 51, 183–194. doi: 10.1026/0012-1924.51.4.183

Schwarz, N., and Bless, H. (1992). "Constructing reality and its alternatives: assimilation and contrast effects in social judgment," in *The Construction of Social Judgments*. eds. L. L. Martin and A. Tesser (Hillsdale, New Jersey: Lawrence Erlbaum Associates), 217–245.

Shavelson, R. J., Hubner, J. J., and Stanton, G. C. (1976). Self-concept: validation of construct interpretations. *Rev. Educ. Res.* 46:407. doi: 10.2307/1170010

Szumski, G., and Karwowski, M. (2015). Emotional and social integration and the big-fish-little-pond effect among students with and without disabilities. *Learn. Individ. Differ.* 43, 63–74. doi: 10.1016/j.lindif.2015.08.037

Tauer, J. M., and Harackiewicz, J. M. (2004). The effects of cooperation and competition on intrinsic motivation and performance. *J. Pers. Soc. Psychol.* 86, 849–861. doi: 10.1037/0022-3514.86.6.849

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., and Imai, K. (2014). Mediation: R package for causal mediation analysis. *J. Stat. Softw.* 59, 1–38. doi: 10.18637/jss. v059.i05

Trautwein, U., Gerlach, E., and Lüdtke, O. (2008). Athletic classmates, physical self-concept, and free-time physical activity: a longitudinal study of frame of reference effects. *J. Educ. Psychol.* 100, 988–1001. doi: 10.1037/0022-0663.100.4.988

Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., and Baumert, J. (2006). Tracking, grading, and student motivation: using group composition and status to predict self-concept and interest in ninth-grade mathematics. *J. Educ. Psychol.* 98, 788–806. doi: 10.1037/0022-0663.98.4.788

Trautwein, U., Lüdtke, O., Marsh, H. W., and Nagy, G. (2009). Within-school social comparison: how students perceive the standing of their class predicts academic self-concept. *J. Educ. Psychol.* 101, 853–866. doi: 10.1037/a0016306

van Kraayenoord, C. E., and Schneider, W. E. (1999). Reading achievement, metacognition, reading self-concept and interest: a study of German students in grades 3 and 4. *Eur. J. Psychol. Educ.* 14, 305–324. doi: 10.1007/BF03173117

Zell, E., and Alicke, M. D. (2009). Contextual neglect, self-evaluation, and the frog-pond effect. *J. Pers. Soc. Psychol.* 97, 467–482. doi: 10.1037/a0015453

Zell, E., and Alicke, M. D. (2010). Comparisons over time: temporal trajectories, social comparison, and self-evaluation. *Eur. J. Soc. Psychol.* 40, 375–382. doi: 10.1002/ejsp.737

Zell, E., Strickhouser, J. E., and Alicke, M. D. (2017). Local dominance effects on self-evaluations and intrinsic motivation. *Self Identity* 16, 629–644. doi: 10.1080/15298868.2017. 1295100