



# Contrast and Assimilation Effects on Self-Evaluation of Performance and Task Interest in a Sample of Elementary School Children

Jannis Bosch\* and Jürgen Wilbert

Department of Inclusive Education, University of Potsdam, Potsdam, Germany

Social comparison processes and the social position within a school class already play a major role in performance evaluation as early as in elementary school. The influence of contrast and assimilation effects on self-evaluation of performance as well as task interest has been widely researched in observational studies under the labels big-fish-little-pond and basking-in-reflected-glory effect. This study examined the influence of similar contrast and assimilation effects in an experimental paradigm. Fifth and sixth grade students ( $n = 230$ ) completed a computer-based learning task during which they received social comparative feedback based on  $2 \times 2$  experimentally manipulated feedback conditions: social position (high vs. low) and peer performance (high vs. low). Results show a more positive development of task interest and self-evaluation of performance in both the high social position and the high peer performance condition. When applied to the school setting, results of this study suggest that students who already perform well in comparison to their peer group are also the ones who profit most from social comparative feedback, given that they are the ones who usually receive the corresponding positive performance feedback.

**Keywords:** social comparison, feedback, self-evaluation, interest, contrast effect

## OPEN ACCESS

### Edited by:

Monica Pellerone,  
Kore University of Enna, Italy

### Reviewed by:

Isabelle Schmidt,  
GESIS Leibniz Institute for the Social  
Sciences, Germany  
Teresa Pozo-Rico,  
University of Alicante, Spain

### \*Correspondence:

Jannis Bosch  
jbosch@uni-potsdam.de

### Specialty section:

This article was submitted to  
Educational Psychology,  
a section of the journal  
Frontiers in Education

**Received:** 18 June 2019

**Accepted:** 30 December 2019

**Published:** 12 February 2020

### Citation:

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

## INTRODUCTION

School is a place where children spend a substantial amount of time in order to learn and develop their academic skills. The social nature of school enables the use of a wide array of group-based learning activities that can have several positive effects on the academic learning process (Springer et al., 1999). However, besides actual academic learning, school also is a place where every student's learning progress and knowledge is constantly evaluated. Performance evaluation is usually carried out by comparing the performance of one student against a reference norm. According to Rheinberg (1980), reference norms can be either criterial (i.e., comparison of performance with an external criterion), ipsative (i.e., comparison of performance with an earlier performance of the same person), or social (i.e., comparison of performance with another person or a group of persons). In school, the social setting almost unavoidably results in a focus on social comparison processes and social comparative performance feedback in the evaluation of each student's academic performance. Accordingly, social-comparative information also plays a major role in each student's self-evaluation of performance (Dijkstra et al., 2008). To differentiate between single instances of self-evaluation and broader self-related constructs (such as the self-concept), the

term self-evaluation will be used within this paper when referring to self-related judgments about a specific task performance.

In general, positive self-evaluation processes are vital for student's well-being (Ames, 1992) and can also be used to retain a positive self-perception in threatening situations (Wills, 1981; Ross and Wilson, 2003). While self-evaluation and more global self-related constructs are not the same, self-evaluation does play a major role in the development of several of those more global self-related judgments, such as self-concept and interest. The self-concept, on the one hand, describes a person's perceptions about themselves in a specific domain. Depending on the context and available information, these perceptions are based on past self-evaluations in the corresponding domain. The academic self-concept, for example, is based on the sum of currently available past self-evaluations of performance in the academic domain (Marsh and Shavelson, 1985). These past self-evaluations are in turn influenced by environmental reinforcements (e.g., external feedback) and by evaluations of significant others (Shavelson et al., 1976). Interest, on the other hand, is defined as the motivational orientation of a person toward a specific object, domain, or area of knowledge (Schiefele, 1992) and can be described as domain-specific intrinsic motivation. Several theories, such as the Self-Determination Theory (SDT; Deci and Ryan, 1985), have suggested that development of interest in a domain is strongly facilitated by positive competence experiences (i.e., positive self-evaluations of performance) in that same domain. Further, academic self-concept and interest have been shown to positively influence each other's development (Marsh et al., 2005; Denissen et al., 2007) and both positively affect future academic performance (Guay et al., 2003; Valentine et al., 2004; Huang, 2011). In fact, some authors have even conceptualized the constructs described as self-concept and interest in the previous section as the respective cognitive (i.e., competence) and affective (i.e., value) components of the self-concept (e.g., Wigfield et al., 1997; Marsh and Ayotte, 2003), further underlining the close connection between both. Additionally, the association between academic self-concept and interest increases with age during school years (Wigfield et al., 1997; Marsh and Ayotte, 2003; Denissen et al., 2007). In summary, academic interest and self-concept are closely related to each other as well as to self-evaluation processes.

Hence, the effects of social comparison processes on self-evaluation (i.e., single instances of self-related judgments), as well as successive effects on broader self-related constructs such as self-concept and interest, are worth taking a closer look at. Generally, the effects of social comparison processes on self-evaluation of performance and the self-concept are increasing during school years (Ruble et al., 1980; Keil et al., 1990) while the academic self-concept itself is declining (Wouters et al., 2012). This could at least partly be explained by a longer exposure to social comparative evaluation structures prevalent in most educational systems (Frey and Ruble, 1985) and the corresponding negative effects on the academic self-concept, particularly in students with poor academic performance. The results of studies investigating the choice of direction of comparison (i.e., comparison with a better or worse performing peer) on self-evaluation and self-concept indicate that choosing

an upward social comparison target (i.e., a better performing peer) can be beneficial, at least in some cases (e.g., Collins, 1996; Huguet et al., 2001). However, there is also evidence suggesting that continuously being at the bottom of the performance range within one's reference group has a negative effect on self-evaluation of performance and, in the long run, also on the academic self-concept and interest (Ames, 1992). Hence, while there is no general detrimental effect of comparison with a better performing peer, a low position within the reference group, and therefore an overabundance of upward comparison targets, appears to have a negative effect on self-evaluation processes.

The nature of the relationship between reference group and self-evaluation of performance can be explained by the Inclusion/Exclusion Model (IEM; Bless and Schwarz, 2010). According to the IEM, the evaluation of a stimulus (e.g., one's own performance in a task) is based on both the stimulus itself and a standard of comparison (e.g., the performance of the rest of the class). Within the IEM two possible effects that can occur when evaluating a stimulus against a standard of comparison are considered: Firstly, a *contrast* effect occurs when the standard of comparison is used as a reference against which the stimulus is compared, i.e., when the student's performance is contrasted against the performance of the rest of the class. Secondly, an *assimilation* effect occurs when properties of the standard of comparison are transferred over to the stimulus, i.e., when the performance of a student in a high performing class is evaluated more positively because of his affiliation with the high performing class.

*Contrast* and *assimilation* effects in academic self-evaluation are the respective assumed bases of popular *big-fish-little-pond* (BFLP) and *basking-in-reflected-glory* (BIRG) effects. BFLP and BIRG effects were first described by Marsh and Parker (1984) and have since then been investigated in a wide array of studies. The BFLP effect can be interpreted as the application of a contrast effect to the school setting. It is based on the observation that students in lower performing classes show a more positive self-concept in comparison to similar performing students in better performing classes. The theory suggests that the difference between a student's performance and the average performance of a salient reference group shapes that student's self-evaluation of performance and therefore, in the long run, also the development of the self-concept. Hence, a negative relationship between a student's academic self-concept and the average performance of that student's class can be expected (Marsh and Parker, 1984). The existence of the BFLP effect has been confirmed in various cross-sectional studies covering different age groups from fifth grade until senior year (e.g., Zeidner and Schleyer, 1999; Köller et al., 2000; Nagengast and Marsh, 2011), as well as in various culturally diverse countries (e.g., Seaton et al., 2009; Wang, 2015). Expectations according to the BIRG effect, on the other hand, are contrary. Because they are aware of the higher perceived standing of their reference group, students in better performing classes are expected to judge their own competence more positively than equally performing students in worse performing classes. Hence, the BIRG effect can be interpreted as the application of an assimilation effect to the school setting. The existence of the BIRG effect could also be confirmed empirically from fifth

grade until senior year (Marsh et al., 2000; Trautwein et al., 2009; Preckel and Brüll, 2010). However, since the BIRG effect is usually smaller than the BFLP effect, the net effect of average class performance on the academic self-concept is still expected to be negative (Marsh et al., 2000). Both the BFLP and the BIRG effect are thought to be based on changes in self-evaluation of performance when confronting students with different standards of comparison (Marsh, 1987).

While BFLP effects on academic self-concept have been extensively studied, similar contrast effects of student performance could also be shown, albeit considerably less well-researched, to affect the development of academic interest (Köller et al., 2000; Trautwein et al., 2006; Schurtz et al., 2014). Further, Trautwein et al. (2006) also tested for assimilation effects similar to the BIRG effect on task interest in 9th grade students by comparing different track levels. Though their results do not show evidence for an assimilation effect on task interest, they also did not find evidence for the assimilation effect on academic self-concept, rendering those results somewhat fragile.

Hence, there is a wide array of studies with a correlational approach showing (a) cross-sectional and longitudinal associations between self-concept and interest in a specific domain, and (b) the associations between average performance of the school class or reference group in a specific domain and an individual's self-concept and interest in that same domain. While (a) can be interpreted as an argument supporting a common mechanism strongly influencing the development of both self-concept and interest over time, (b) can be interpreted as evidence that this mechanism is somewhat dependent on social comparison processes.

## Research Question and Hypotheses

Past research in the field mainly focused on the description of the presented phenomena in non-experimental research designs. Within this paper, we want to investigate whether the mechanisms behind popular BFLP and BIRG effects on academic self-concept, namely changes in self-evaluation based on different standards of reference, are also responsible for contrast and assimilation effects observed in academic interest (e.g., Köller et al., 2000). Therefore, this study sets out to close a research gap in the field by closely investigating the mechanism behind contrast and assimilation effects on task interest that have been found in several observational studies presented earlier (e.g., Köller et al., 2000; Trautwein et al., 2006; Schurtz et al., 2014). Self-evaluation of performance has previously been suggested as a factor connecting self-concept and interest (Denissen et al., 2007), and it has also been described as a major contributor to interest development by Deci and Ryan (1985) in their SDT. To be able to have a close look at these mechanisms, the present study was aimed at experimentally manipulating feedback of personal performance in a social context (from now on referred to as *social position*) as well as the criterial performance of the reference group (from now on referred to as *peer performance*). To the author's knowledge, there are only two studies examining the direct effects of experimentally manipulated social position in an academic learning task (Pohlmann and Möller, 2006; Bosch and Wilbert, 2017). However, none of them investigated school

age children. Pohlmann and Möller (2006) gave manipulated social comparative feedback to university students on two different academic tasks, one concerning word analogies and one concerning figure analogies. They were able to show that a higher social position was clearly associated with a more positive self-evaluation of performance ( $d = 1.89$  for word analogies and  $d = 1.92$  for figure analogies). Results concerning interest were less conclusive. While they could show a similar contrast effect in the word analogies task ( $d = 0.41$ ), no contrast effect could be shown in the figure analogies task ( $d = 0.00$ ). Assimilation effects were not within the scope of their study. Bosch and Wilbert (2017) investigated both contrast and assimilation effects in an academic learning task in a sample of university students. They only included task interest and task performance as dependent variables, but no measure of self-evaluation or self-concept. Results showed a contrast effect ( $d = 0.22$ ) as well as an *assimilation* effect ( $d = 0.20$ ) on task interest. However, the latter only showed a trend toward significance<sup>1</sup>. There were no contrast or assimilation effects on task performance. The lack of effects on task performance could be due to the very short nature of the experiment (roughly 30 min). Short-term negative effects of social comparative performance feedback on actual task performance are suggested to be based on an "inwards" focus that takes away some of the attention from the task but is independent of feedback valence (Kluger and DeNisi, 1996).

Hypothesis 1: Hence, based on previous empirical results and theoretical considerations, we expect a higher social position (compared to a lower social position) within a reference group as well as a higher peer performance (compared to a lower peer performance) to be associated with a more positive self-evaluation of performance.

Hypothesis 2: We also expect similar positive effects of social position and peer performance on the development of task interest. We expect these effects to be smaller than those on self-evaluation of performance.

Hypothesis 3: We do not expect similar effects in relation to the development of task performance, because of the short duration of the experiment.

Further, this study sets out to replicate the previously presented positive associations between self-concept, interest, and performance in elementary school children.

## METHODS

### Participants

The sample consisted of 230 elementary school students (122 girls, 108 boys) from 16 classes in four different schools in the Brandenburg region in Germany. Ninety-seven participants were in fifth grade and 133 were in sixth grade. The age of tested participants ranged from 10–13 years; the average student was 11.31 years ( $SD = 0.74$ ) old.

<sup>1</sup>Effect sizes of Bosch and Wilbert (2017) were calculated using the procedure suggested by Morris (2008) as  $d_{ppc2}$  on raw data provided by the authors. 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).

## Learning Task

The learning task used in this study was an adapted version of the “flag game,” a computer-based learning game used in a previous study with university students. In the “flag game” participants are requested to learn pairs of flags and corresponding country outlines. The adaptation for elementary school children included a substantial reduction of the total items (number of pairs) to be learned and a rework of the instruction in a language comprehensible by fifth and sixth graders. In order to make sure the adaptation of the task was appropriate, a preliminary study was carried out with an independent sample ( $N = 52$ ). A small subsample of these students ( $N = 6$ ) was also interviewed afterwards to identify potential obscurities for the target group. All students interviewed were able to follow the instructions and understood how to perform the task. Further, none of the students reported distress by the feedback provided.

Similar to the original (Bosch and Wilbert, 2017), each run of the adapted version of the “flag game” consisted of two phases: a learning and a performance phase. During the first learning phase, participants were presented with a map of the Northern half of the African continent as well as the outlines of every country shown. For each of the 17 learning trials, the outlines of one of these countries were highlighted, accompanied by the corresponding national flag. Each combination was presented for a maximum of 10 s and participants could shorten the presentation by pressing the space key. After finishing the first learning phase, the first performance phase was started. The performance phase also consisted of 17 trials and during each, one of the 17 previously learned flags was presented with five different highlighted country outlines (i.e., the target country and four distractor countries). After the first run, performance feedback was given based on the respective experimental condition. Then, a second run was started with another learning and performance phase. However, this time 17 countries from the Southern half of the African continent were presented. After the second run, no feedback was given. To avoid sequencing effects between conditions, trial sequences were pseudo-randomized for each run, and the resulting sequences were then used for all participants in all experimental conditions. Randomization was carried out separately for each phase and run, so each phase had a unique random sequence. Distractor items were taken from each phases’ item pool, and balanced and pseudo-randomized separately for each performance phase. Hence, every participant was presented with the exact same trial sequences and distractor items.

## Experimental Feedback Conditions

During the instruction, participants were told they would receive feedback with information about their own performance compared to the performance of other elementary school children. However, feedback did not reflect their real performance, but was based on respective experimental conditions. To test for potential contrast effects, social position (SP; high vs. low social position) was altered and to test for potential assimilation effects, peer performance (PP; high vs. low peer performance) was altered. Therefore, the study was based on  $2 \times 2$  experimental conditions and each participant was

randomly assigned to one of four groups: 58 participants were in the high social position/high peer performance condition (SP+/PP+), 57 participants were in the low social position/high peer performance condition (SP-/PP+), 59 participants were in the high social position/low peer performance condition (SP+/PP-), and 56 participants were in the low social position/low peer performance condition (SP-/PP-).

During the instruction period, participants were told that feedback was given on a scale from 0 to 100 based on both correctness and speed of their responses. Response speed was included to prevent participants from predicting their own performance. During the feedback phase, participants in high peer performance conditions were shown a mean score of 64 points for other children from their elementary school, while participants in low peer position conditions were shown a mean score of 34 points for other children from their elementary school. Participants in high social position conditions received feedback indicating they scored 15 points above average elementary school children from their school. Hence, participants in the high social position/high peer performance condition were told they scored 79 points, while participants in the high social position/low peer performance condition were told they scored 49 points. Participants in low social position conditions received feedback indicating they scored 15 points below average. Therefore, participants in the low social position/high peer performance condition were told they scored 49 points, while participants in the low social position/low peer performance condition were told they scored 19 points. Feedback was given after the first run of the flag game. Participants were first shown a slide with their own performance. On a second slide they were then shown how their fellow elementary school children allegedly performed on the task.

## Procedure

Prior to gathering the data, consent was obtained from each schools’ head of school as well as from all teachers involved. Roughly 1 week prior to testing, a consent form to be signed by the parents was handed out to each of the students of participating classes. Only students who returned the signed consent form could participate in the study. On testing day, students were told that they were about to play a computer-based learning game and informed that they could terminate their participation at any given time. They were also briefed on the exact procedure of the study. Then, participating students were randomly assigned to groups of four people. Each group was individually taken out of regular class and led into a separate room. Once in the room, each student was randomly allocated one of four visually shielded laptops. When seated, participants were asked to fill out a short questionnaire regarding age, gender and grade level and a questionnaire about their self-concept in learning tasks. After all students finished filling out the questionnaires, they were asked to start the computer-based instruction of the “flag game.” After finishing the instruction, they were asked to fill out the first task interest questionnaire. They were then instructed to start the computer-based learning game. After receiving experimentally manipulated feedback on their performance during the first performance

phase, participants' self-evaluation of performance was assessed. Then the second run was started. No feedback was given after the second run. After all participants finished the computer-based learning task, they were asked to fill out the task interest questionnaire a second time to investigate potential changes in interest. In the end, participants were asked whether they noticed anything unusual or had any suspicions about the performance feedback they received during the task. These questions were used as a manipulation check and were independently analyzed by two researchers. Neither found any hints of exposure of the experimental manipulation. After all students of a class finished the task, both the true nature of the study and the experimental manipulation were disclosed. Further, students were told that everyone did very well on the task.

## Instruments

Because all participants were immediately notified of any missing items, there was no missing data in our dataset.

### Task-Specific Self-Concept

The self-concept scale consists of three statements about the participant's self-rated ability to resolve learning tasks: One positive ("Usually I have no trouble with learning tasks") and two negative (e.g., "Dealing with learning tasks isn't one of my strengths") statements. The questionnaire is an adapted version of a self-concept questionnaire used by Bosch and Wilbert (2017). Participants rated the items immediately after the instruction (T1) on a Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree). Self-concept was used as a control variable. Negative items were reversed, and a mean self-concept score was calculated. Cronbach's alpha of the scale was  $\alpha = 0.63$ .

### Task-Specific Interest

The interest scale consists of four items concerning interest in the task (e.g., "I like learning tasks like this one"). The items are adapted versions of an interest scale from a German motivation questionnaire (Rheinberg et al., 2001). Each item was rated on a Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree), after the instruction of the learning task (T1 interest) and once after finishing the learning task (T2 interest). T1 interest had a Cronbach's alpha of  $\alpha = 0.86$ , while T2 interest had a Cronbach's alpha of  $\alpha = 0.87$ .

### Task Performance

Performance was measured as percentage of correct answers in both runs of the "flag game." T1 performance was the proportion of correctly answered multiple-choice items in the first performance phase, while T2 performance was the ratio of correctly answered items in the second performance phase. Hence, obtainable scores ranged from 0 to 1.

### Self-Evaluation of Performance

Self-evaluation of performance was assessed directly after participants finished the first performance phase of the "flag game" and received the experimentally manipulated performance feedback. In contrast to the task-specific self-concept assessed prior to performing the task, self-evaluation of performance was conceptualized as a self-related judgment about a single instance

of task performance. Hence, participants were asked to evaluate their own performance shortly after the first run of the "flag game" on a Likert-type scale ranging from 1 (very bad) to 5 (very good).

## Statistical Analyses

Associations between self-concept, task interest, and task performance were investigated using correlation analyses.

Experimental conditions were effect-coded for all subsequent analyses ( $-0.5$  for low social position and low peer performance;  $+0.5$  for high social position and high peer performance, respectively). This means the *b*-weights reported reflect the mean difference between experimental conditions (in unifactorial cases).

Intraclass correlations did not show a substantial amount of variance explained by class membership for any of our dependent variables ( $ICC_{class\ self-evaluation} = 0.00$ ,  $ICC_{class\ task\ interest} = 0.02$ ,  $ICC_{class\ task\ performance} = 0.03$ ). These results suggest that students from different classes tested in this study were relatively homogeneous in relation to the tested dependent variables. Hence, school class membership was not included in further analyses.

To test the earlier presented hypothesis concerning the effects of *social position* and *peer performance* in social comparative performance feedback on self-evaluation of performance, we used multiple regression models with self-evaluation of performance as criterion. Predictors were subsequently added to the model based on the respective hypotheses tested. Model 1 was the null-model and therefore only included the intercept. For model 2, we added *social position* as predictor. Model 3 further included *peer performance* and the interaction between *social position* and *peer performance*. In order to control for potential interactions with initial *self-concept*, we added centered *self-concept* and all interaction terms with previously added predictors to model 4. Adjusted  $R^2$  of each model was reported and *F*-tests were calculated for subsequent models to determine whether each set of predictors significantly increased total explained variance.

In order to test for similar effects of *social position* and *peer performance* on the development of task interest and task performance during the computer-based learning task, we used linear mixed models (fixed and random effects), as suggested for repeated measures regression analyses by Everitt and Hothorn (2011). Linear mixed models were calculated with the *nlme* package for R (Pinheiro et al., 2019). Once again, predictors were subsequently added to the model. Models used to predict development of task interest and performance only differed in the respective criterion, predictors were added in a similar fashion. Model 1 only included *measurement time* as predictor. Contrasts for *measurement time* were coded as 0 for T1 and 1 for T2 for both the interest and the performance model. Therefore, the intercept reflects the mean of measurements at T1. Model 2 included both *measurement time* and *social position*, as well as the interaction between *measurement time* and *social position*. Model 3 further included *peer performance* and all interaction terms. Because regular  $R^2$  cannot be calculated for linear mixed models,  $R^2\ Beta$  from the *r2glmm* package for R (Jaeger, 2017) was

**TABLE 1 |** Intercorrelations for all observed variables ( $N = 230$ ).

|                    | 1                 | 2       | 3        | 4       | 5                 | 6       | 7    |
|--------------------|-------------------|---------|----------|---------|-------------------|---------|------|
| 1. Self-concept    | 1                 |         |          |         |                   |         |      |
| 2. Self-evaluation | -0.02             | 1       |          |         |                   |         |      |
| 3. T1 interest     | 0.13 <sup>†</sup> | 0.02    | 1        |         |                   |         |      |
| 4. T2 interest     | 0.10              | 0.16*   | 0.79***  | 1       |                   |         |      |
| 5. Interest diff.  | -0.04             | 0.22*** | -0.25*** | 0.40*** | 1                 |         |      |
| 6. T1 performance  | 0.15*             | 0.03    | 0.04     | 0.09    | 0.09              | 1       |      |
| 7. T2 performance  | 0.14*             | -0.01   | -0.03    | 0.05    | 0.11 <sup>†</sup> | 0.43*** | 1    |
| 8. Perf. diff.     | -0.01             | -0.04   | -0.06    | -0.04   | 0.02              | -0.52   | 0.54 |

\*\*\* $p < 0.001$ , \* $p < 0.05$ , <sup>†</sup> $p < 0.10$ ; Interest diff., difference between T1 and T2 interest; Perf. diff., difference between T1 and T2 performance.

used to determine variance explained by fixed effects as suggested by Jaeger et al. (2017). Further, the corrected Akaike Information Criterion ( $AIC_c$ ) is reported as an additional measure to compare presented models (smaller values resemble a better relative model fit; Hurvich and Tsai, 1989).

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

## RESULTS

**Table 1** shows intercorrelations between all observed variables. As can be seen, self-concept shows a moderately positive correlation with both T1 and T2 interest, even though only the former reaches statistical significance. Further, it shows a significant positive correlation of a similar magnitude with T1 and T2 performance. The expected positive associations between interest and performance could not be found in the present sample. Further, there was no association between initial self-concept and self-evaluation of performance. Hence, results support the hypothesized associations between self-concept and interest and self-concept and performance, but not between interest and performance.

**Table 2** shows means and standard deviations for all relevant study variables both for the complete sample and separately for each experimental condition. To test for potential group differences despite randomization, one-factorial ANOVAs with experimental condition as independent variable were conducted for all variables assessed at T1, before the intervention took place. Results did not show any significant group differences in self-concept,  $F_{(3,226)} = 1.13$ ,  $p = 0.34$ , T1 interest,  $F_{(3,226)} = 1.38$ ,  $p = 0.25$ , or T1 performance,  $F_{(3,226)} = 0.32$ ,  $p = 0.81$ , suggesting randomization led to similar distributions across experimental groups.

### Hypothesis 1: Effects of Social Position and Peer Performance on Self-Evaluation of Performance

**Table 3** shows model fit and parameter estimates for all four linear regression models predicting self-evaluation of performance. Model 2 containing *social position* as a predictor

**TABLE 2 |** Descriptive statistics for all observed variables separately for each experimental condition ( $N = 230$ ).

|                                       | SP +<br>PP + | SP -<br>PP + | SP +<br>PP - | SP -<br>PP - | All<br>participants |
|---------------------------------------|--------------|--------------|--------------|--------------|---------------------|
| <b>Self-concept</b>                   |              |              |              |              |                     |
| T1                                    | 5.29 (1.29)  | 4.99 (1.36)  | 5.16 (1.33)  | 5.41 (1.18)  | 5.21 (1.29)         |
| <b>Self-evaluation of performance</b> |              |              |              |              |                     |
| T2                                    | 3.95 (0.69)  | 2.86 (0.69)  | 3.22 (0.67)  | 2.23 (0.89)  | 3.07 (0.96)         |
| <b>Interest</b>                       |              |              |              |              |                     |
| T1                                    | 4.99 (1.40)  | 4.77 (1.56)  | 5.11 (1.54)  | 5.19 (1.23)  | 5.02 (1.44)         |
| T2                                    | 5.19 (1.24)  | 4.53 (1.64)  | 4.90 (1.57)  | 4.58 (1.55)  | 4.80 (1.52)         |
| Difference                            | 0.19 (1.06)  | -0.24 (0.73) | -0.21 (0.93) | -0.62 (0.95) | -0.21 (0.96)        |
| <b>Performance</b>                    |              |              |              |              |                     |
| T1                                    | 0.41 (0.21)  | 0.39 (0.14)  | 0.41 (0.17)  | 0.39 (0.17)  | 0.40 (0.17)         |
| T2                                    | 0.42 (0.18)  | 0.42 (0.18)  | 0.43 (0.17)  | 0.41 (0.17)  | 0.42 (0.18)         |
| Difference                            | 0.01 (0.20)  | 0.03 (0.16)  | 0.02 (0.17)  | 0.03 (0.21)  | 0.02 (0.18)         |

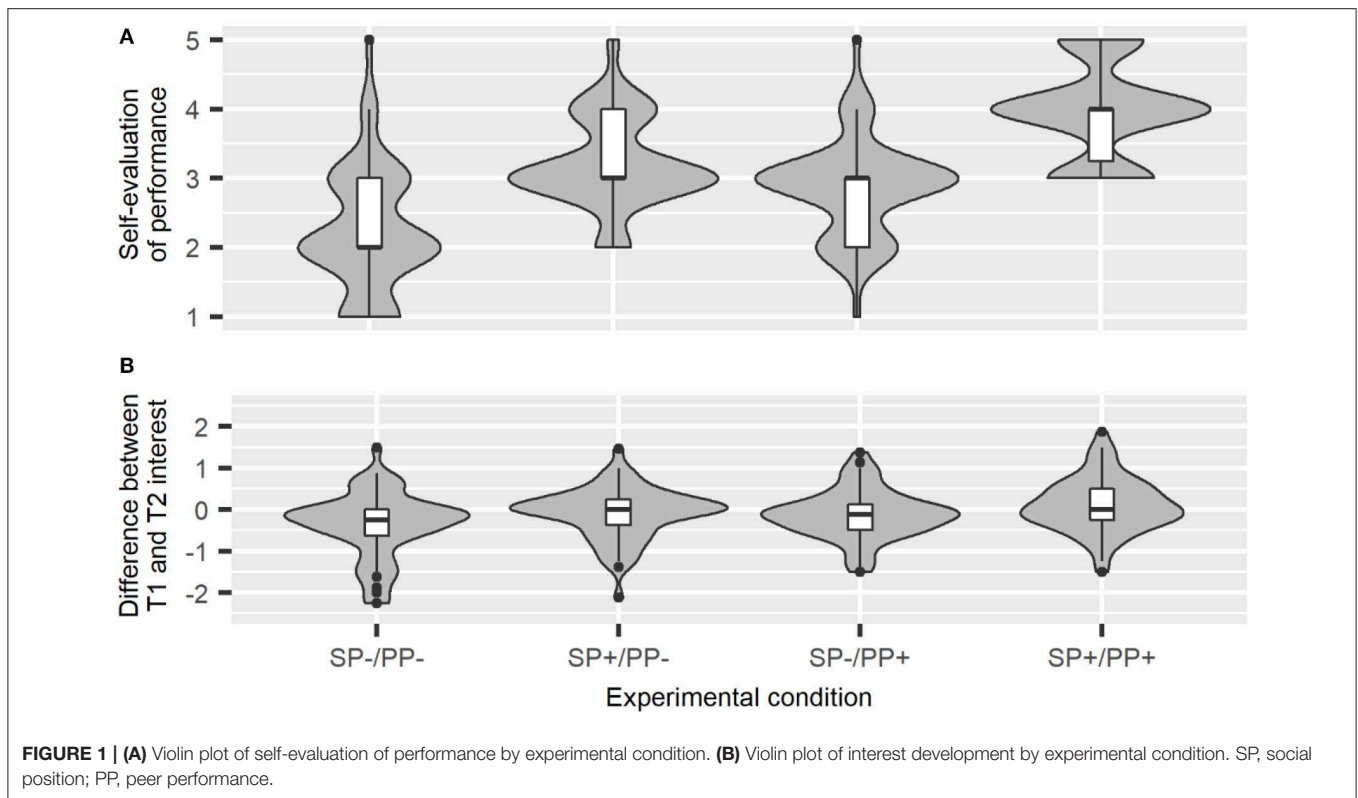
SP, social position; PP, peer performance; Performance variables reflect percentage of correct answers.

**TABLE 3 |** Unstandardized estimates for linear regression models predicting self-evaluation of performance after the feedback intervention ( $N = 230$ ).

| Parameter            | Unstandardized estimate b (SE) |                 |                 |                 |
|----------------------|--------------------------------|-----------------|-----------------|-----------------|
|                      | Model 1                        | Model 2         | Model 3         | Model 4         |
| Intercept            | 3.07*** (0.063)                | 3.07*** (0.054) | 3.07*** (0.049) | 3.07*** (0.049) |
| SP                   |                                | 1.03*** (0.107) | 1.04*** (0.098) | 1.02*** (0.099) |
| PP                   |                                |                 | 0.68*** (0.098) | 0.67*** (0.099) |
| SC                   |                                |                 |                 | -0.01 (0.038)   |
| SP × PP              |                                |                 | 0.10 (0.195)    | 0.11 (0.197)    |
| SP × SC              |                                |                 |                 | 0.05 (0.077)    |
| PP × SC              |                                |                 |                 | 0.14 (0.077)    |
| SP × PP × SC         |                                |                 |                 | -0.01 (0.154)   |
| <b>Model indices</b> |                                |                 |                 |                 |
| df                   | 229                            | 228             | 226             | 222             |
| Adj. R <sup>2</sup>  | 0.000                          | 0.286           | 0.407           | 0.406           |

\*\*\* $p < 0.001$ , SP, social position (-0.5/+0.5); PP, peer performance (-0.5/+0.5); SC, self-concept (centered); df, degrees of freedom; Adj. R<sup>2</sup>, Adjusted R<sup>2</sup>.

explained significantly more variance compared to the intercept-only model 1,  $\Delta R^2 = 0.286$ ,  $F_{(1,228)} = 111.57$ ,  $p < 0.001$ . Addition of *peer performance* and the interaction between *social position* and *peer performance* in model 3 increased variance explained,  $\Delta R^2 = 0.121$ ,  $F_{(2,226)} = 24.23$ ,  $p < 0.001$ , compared to model 2, while addition of centered *self-concept* and all interactions with previous predictors (model 4) did not further increase variance explained,  $\Delta R^2 = -0.001$ ,  $F_{(4,222)} = 0.87$ ,  $p = 0.485$ . A closer look at the regression weights in models 3 and 4 shows that *social position* has a stronger effect on self-evaluation of performance than *peer performance*. **Figure 1A** shows mean differences in self-evaluation of performance between experimental conditions as well as the corresponding distributions. Hence, both *social position* and *peer performance* showed positive regression weights (i.e., a higher social position and a higher peer performance



both lead to a more positive self-evaluation of performance), supporting hypothesis 1. Initial self-concept did not change this pattern, nor did it affect the self-evaluation independent of experimental condition.

## Hypothesis 2: Effects of Social Position and Peer Performance on Task Interest

Table 4 shows model and parameter estimates for linear mixed models predicting task interest. Model 1, only containing *measurement time* as predictor, explained significantly more variance than the intercept-only model ( $\Delta R^2\text{Beta} = 0.008$ , *Likelihood Ratio* = 11.09,  $p < 0.001$ ). The addition of *social position* and the interaction between *measurement time* and *social position* in model 2 improved variance explained ( $\Delta R^2\text{Beta} = 0.023$ , *Likelihood Ratio* = 13.37,  $p < 0.01$ ). While addition of *peer performance* and its interactions (model 3) with the previous predictors further improved variance explained, the change in model fit barely failed to reach statistical significance ( $\Delta R^2\text{Beta} = 0.014$ , *Likelihood Ratio* = 11.23,  $p < 0.05$ ). In conclusion, participants showed a general decline in interest from T1 to T2. However, these changes were affected by the experimental manipulation of *social position* and *peer performance*. As suggested in hypothesis 2, respective high *social position* and *peer performance* conditions showed a positive influence on the development of task interest compared to respective low *social position* and *peer performance* conditions. When comparing the regression weights for the

interaction between *social position* and *measurement time*, and *peer performance* and *measurement time* in model 3, one can see that *social position* has a slightly stronger effect on interest development. Differences in interest development between experimental conditions and corresponding distributions can be seen in Figure 1B. The only condition with an absolute increase in interest from T1 to T2 was the high *social position*/high *peer performance* group, while all other conditions showed an absolute decline in interest (see Table 2).

## Hypothesis 3: Effects of Social Position and Peer Performance on Task Performance

Table 5 shows model and parameter estimates for linear mixed models predicting task performance. Model 1 contained only *measurement time* as predictor and significantly increased model fit compared to the intercept-only model ( $\Delta R^2\text{Beta} = 0.005$ , *Likelihood Ratio* = 3.88,  $p < 0.05$ ). The addition of *social position* and its interaction with *measurement time* in model 2 ( $\Delta R^2\text{Beta} = 0.003$ , *Likelihood Ratio* = 0.93,  $p = 0.63$ ), as well as the further addition of *peer performance* and its interactions in model 3 ( $\Delta R^2\text{Beta} = 0.000$ , *Likelihood Ratio* = 0.12,  $p = 0.99$ ) did not lead to increases in explained variance. Hence, while there was a small overall increase in performance from T1 to T2, there were no differences in performance development between experimental groups. That suggests the experimental manipulation did not influence participants' performance.

**TABLE 4** | Fixed effects for mixed models predicting interest ( $N = 230$ ).

| Parameter            | Unstandardized estimate b (SE) |                  |                  |
|----------------------|--------------------------------|------------------|------------------|
|                      | Model 1                        | Model 2          | Model 3          |
| Intercept            | 5.02*** (0.098)                | 5.02*** (0.097)  | 5.02*** (0.097)  |
| <b>Level 1</b>       |                                |                  |                  |
| MT                   | -0.21*** (0.063)               | -0.22*** (0.062) | -0.22*** (0.061) |
| <b>Level 2</b>       |                                |                  |                  |
| SP                   |                                | 0.07 (0.194)     | 0.07 (0.194)     |
| PP                   |                                |                  | -0.27 (0.194)    |
| <b>Interactions</b>  |                                |                  |                  |
| MT × SP              |                                | 0.42** (0.124)   | 0.42*** (0.122)  |
| MT × PP              |                                |                  | 0.39** (0.122)   |
| SP × PP              |                                |                  | 0.30 (0.389)     |
| MT × SP × PP         |                                |                  | 0.02 (0.244)     |
| <b>Model indices</b> |                                |                  |                  |
| $AIC_c$              | 1448                           | 1439             | 1436             |
| $df$                 | 4                              | 6                | 10               |
| $R^2$ Beta           | 0.008                          | 0.031            | 0.045            |

\*\*\* $p < 0.001$ , \*\* $p < 0.01$ , MT, measurement time (0/1); SP, social position (-0.5/+0.5); PP, peer performance (-0.5/+0.5);  $AIC_c$ , corrected Akaike Information Criterion;  $df$ , degrees of freedom.

**TABLE 5** | Fixed effects for mixed models predicting task performance ( $N = 230$ ).

| Parameter            | Unstandardized estimate b (SE) |                 |                 |
|----------------------|--------------------------------|-----------------|-----------------|
|                      | Model 1                        | Model 2         | Model 3         |
| Intercept            | 0.40*** (0.012)                | 0.40*** (0.012) | 0.40*** (0.012) |
| <b>Level 1</b>       |                                |                 |                 |
| MT                   | 0.02* (0.012)                  | 0.02* (0.012)   | 0.02† (0.012)   |
| <b>Level 2</b>       |                                |                 |                 |
| SP                   |                                | 0.02 (0.023)    | 0.02 (0.023)    |
| PP                   |                                |                 | 0.00 (0.023)    |
| <b>Interactions</b>  |                                |                 |                 |
| MT × SP              |                                | -0.01 (0.024)   | -0.01 (0.025)   |
| MT × PP              |                                |                 | -0.00 (0.025)   |
| SP × PP              |                                |                 | 0.00 (0.046)    |
| MT × SP × PP         |                                |                 | -0.02 (0.049)   |
| <b>Model indices</b> |                                |                 |                 |
| $AIC_c$              | -347                           | -344            | -336            |
| $df$                 | 4                              | 6               | 10              |
| $R^2$ Beta           | 0.005                          | 0.008           | 0.008           |

\*\*\* $p < 0.001$ , \* $p < 0.05$ , † $p < 0.10$ ; MT, measurement time (0/1); SP, social position (-0.5/+0.5); PP, peer performance (-0.5/+0.5);  $AIC_c$ , corrected Akaike Information Criterion;  $df$ , degrees of freedom.

## DISCUSSION

In this study we investigated whether social comparative performance feedback influences self-evaluation of performance, task interest and task performance of elementary school students in an academic learning task. To investigate the mechanisms suggested to be the basis of popular *BFLP* and *BIRG* effects, namely contrast and assimilation effects on self-evaluation of performance, both social position and peer performance were experimentally manipulated in a learning game with a performance feedback intervention.

The clearest differences between feedback conditions could be shown in relation to self-evaluation of performance after receiving the feedback (see **Tables 2, 3**). On average, participants in the high social position conditions evaluated their own performance much more positively than participants in the low social position conditions. Participants in the high social position conditions scored roughly one point higher on the self-evaluation scale, resulting in a positive self-evaluation of performance on average for these conditions. Participants in the low social position conditions, in contrast, showed a neutral to negative self-evaluation of performance. Group differences between high and low peer performance conditions showed a similar pattern but were a bit smaller. As a result, differences in feedback conditions explained around 40% of variance in self-evaluation of performance, showing a major influence of external feedback on self-evaluation of performance. The strong influence of experimental conditions on self-evaluation of performance could also explain why there was no association between self-concept (assessed before experimental manipulation) and self-evaluation of performance (assessed after experimental manipulation)

within our sample, as could be expected for these conceptually related variables.

Hence, the differences between participants in the high and low social position conditions can be interpreted as evidence for the occurrence of contrast effects, as suggested by the IEM (Schwarz and Bless, 1992; Bless and Schwarz, 2010), when students receive information about their own performance in relation to the performance of their classmates. Similarly, the differences between participants in the high and low peer performance conditions can be interpreted as evidence for the occurrence of assimilation effects when students receive information about a high criterial score of their classmates.

While both social position and peer performance had relatively strong effect sizes on self-evaluation of performance, their influence was considerably less pronounced in the conceptually more distant study variable task interest. Once more, however, participants in the high social position and peer performance conditions showed a pattern of interest development similar to the one observed for self-evaluation of performance. Hence, results suggest a more positive development of interest in the high social position and peer performance conditions, with a slightly larger difference in the social position conditions. These results can be interpreted as evidence for the existence of contrast and assimilation effects of social comparative performance feedback on the development of task interest. Since the pattern is somewhat similar to the one observed for self-evaluation of performance, results also provide indirect evidence for the influence of self-evaluation of performance on the development of task interest. However, in order to develop interest in a task or domain, the need



to feel competent is only one of several influencing factors (Ryan and Deci, 2000; Krapp, 2005). Therefore, it is no surprise that experimental conditions can only explain about 4% of the variance in the development of task interest, suggesting a much looser association with external feedback compared to self-evaluation of performance. The small effect size might also be an explanation why Pohlmann and Möller (2006) found evidence for a similar contrast effect of social comparative performance feedback on task interest in one creativity task, but not in the other.

While the effects of social comparative feedback on task interest in this sample of elementary school children were similar to those observed in a sample of university students, correlations of task interest with self-concept and task performance were much smaller (Bosch and Wilbert, 2017). According to Schurtz et al. (2014), young children have relatively broad interests, and interest only gets more differentiated and domain-specific toward adolescence. Thus, the lack of differentiation in interest might explain the relatively small associations interest shows with self-concept and task performance in this sample compared to a sample of university students.

## Implications

These results show that the students most clearly benefitting from social comparative performance feedback are the ones who finish on top of their class, while those receiving feedback about a lower social position tend to evaluate their own performance more negatively which is also accompanied by a loss of interest. As suggested earlier, that could in turn lead to a further decline of academic achievement (Schiefele, 1992; Möller et al., 2011), especially for students who persistently perform worse than their classmates. Hence, the use of social comparative feedback information might be detrimental to the development of interest for at least part of the students. Taken together with the results of studies on longitudinal development of task interest and self-concept, showing that the social position within one's own class does affect both constructs, results of the study at hand suggest the current practice of assigning grades based on social comparative performance information could seriously hurt the further academic development of students who continuously find themselves on the bottom of the performance range within their class or peer group. Further, the results of this study also provide further evidence that the effects of social comparative performance feedback on a person's self-evaluation is at least part of the mechanism behind BFLP and BIRG effects on task interest.

Other studies could show that a focus on task mastery (Butler, 1992) or the learning process itself (Harks et al., 2014) could be beneficial for all students rather than reinforcing interest only in those who already perform well. Hence, while there is no clear best practice concerning performance evaluation and feedback, additional research comparing different feedback practices would be helpful to ensure feedback mechanisms used in institutional education actually facilitate the learning process of all students rather than reinforcing only those who already perform well.

## Limitations

The experimental nature of this study enabled a closer look at the supposed mechanisms behind popular BFLP and BIRG effects. However, the very linear and controlled environment used in the learning task also lacked some of the dimensions of regular classroom learning, such as cooperation between students and teacher-student interactions. Further, the operationalization of the BIRG effect via different criterial scores does lack several dimensions of real-life educational situations (e.g., being selected into a highly valued group; Marsh et al., 2000). However, while the experimental manipulation of grading within regular classrooms would be helpful from the perspective of a researcher, it is questionable whether the increase in external validity of these results would be a sufficient argument to justify experimental manipulation of real grading for actual students. Additionally, the comparison of equally performing students in peer groups of different skill levels, as suggested by the BFLP effect, is only directly possible between the high social position/low peer performance and the high peer performance/low social position conditions where students received a similar criterial feedback of 49 points. However, the general causal mechanisms of the BFLP effect are still covered by the experimental manipulations of this study.

## DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

## ETHICS STATEMENT

This research was conducted in accordance with APA Ethical Guidelines for Research. It was reviewed and approved by Ministry of Education, Youth, and Sports in Brandenburg. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

The study was planned by both authors. The execution was supervised by JB. The manuscript was mostly written by JB, but several parts were co-written by JW. Contents were continuously discussed by both authors.

## FUNDING

We acknowledge the support of Open Access Publication Fund of Potsdam University.

## ACKNOWLEDGMENTS

Thanks to the library of University of Potsdam for paying the open access fees. We would also like to thank all students who helped carrying out the study.

## REFERENCES

- Ames, C. (1992). Classrooms: goals, structures, and student motivation. *J. Educ. Psychol.* 84, 261–271. doi: 10.1037/0022-0663.84.3.261
- 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
- Butler, R. (1992). What young people want to know when: effects of mastery and ability goals on interest in different kinds of social comparisons. *J. Pers. Soc. Psychol.* 62, 934–943. doi: 10.1037/0022-3514.62.6.934
- Collins, R. L. (1996). For better or worse: the impact of upward social comparison on self-evaluations. *Psychol. Bull.* 119, 51–69. doi: 10.1037/0033-2909.119.1.51
- Deci, E. L., and Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York, NY: Plenum.
- Denissen, J. J., Zaretz, 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
- Everitt, B., and Hothorn, T. (2011). *An Introduction to Applied Multivariate Analysis with R*. New York, NY: Springer.
- Frey, K. S., and Ruble, D. N. (1985). What children say when the teacher is not around: conflicting goals in social comparison and performance assessment in the classroom. *J. Pers. Soc. Psychol.* 48, 550–562. doi: 10.1037/0022-3514.48.3.550
- 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
- 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
- Harks, B., Rakoczy, K., Hattie, J., Besser, M., and Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: the role of feedback's perceived usefulness. *Educ. Psychol.* 34, 269–290. doi: 10.1080/01443410.2013.785384
- Huang, C. (2011). Self-concept and academic achievement: a meta-analysis of longitudinal relations. *J. Sch. Psychol.* 49, 505–528. doi: 10.1016/j.jsp.2011.07.001
- Huguet, P., Dumas, F., Monteil, J. M., and Genestoux, N. (2001). Social comparison choices in the classroom: further evidence for students' upward comparison tendency and its beneficial impact on performance. *Eur. J. Soc. Psychol.* 31, 557–578. doi: 10.1002/ejsp.81
- Hurvich, C. M., and Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika* 76, 297–307. doi: 10.1093/biomet/76.2.297
- Jaeger, B. (2017). *r2glmm: Computes R Squared for Mixed (Multilevel) Models*. (Version R Package Version 0.1.2). Retrieved from: <https://CRAN.R-project.org/package=r2glmm>
- Jaeger, B., Edwards, L. J., Das, K., and Sen, P. K. (2017). An  $R^2$  statistic for fixed effects in the generalized linear mixed model. *J. Appl. Stat.* 44, 1086–1105. doi: 10.1080/02664763.2016.1193725
- Keil, L. J., McClintock, C. G., Kramer, R., and Platow, M. J. (1990). Children's use of social comparison standards in judging performance and their effects on self-evaluation. *Contemp. Educ. Psychol.* 15, 75–91. doi: 10.1016/0361-476X(90)90007-N
- Kluger, A. N., and DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol. Bull.* 119, 254–284. doi: 10.1037/0033-2909.119.2.254
- Köller, O., Schnabel, K. U., and Baumert, J. (2000). Der Einfluß der Leistungsstärke von Schulen auf das fachspezifische Selbstkonzept der Begabung und das Interesse [The impact of the school's average performance on students' domain-specific self-concept and interest]. *Z. Entwicklungspsychol. Pädagog. Psychol.* 32, 70–80. doi: 10.1026//0049-8637.32.2.70
- 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
- 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 Ayotte, V. (2003). Do multiple dimensions of self-concept become more differentiated with age? The differential distinctiveness hypothesis. *J. Educ. Psychol.* 95, 687–706. doi: 10.1037/0022-0663.95.4.687
- 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 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., and Shavelson, R. (1985). Self-concept: its multifaceted, hierarchical structure. *Educ. Psychol.* 20, 107–123. doi: 10.1207/s15326985ep2003\_1
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., and Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: reciprocal effects models of causal ordering. *Child Dev.* 76, 397–416. doi: 10.1111/j.1467-8624.2005.00853.x
- Möller, J., Retelsdorf, J., Köller, O., and Marsh, H. W. (2011). The reciprocal internal/external frame of reference model: an integration of models of relations between academic achievement and self-concept. *Am. Educ. Res. J.* 48, 1315–1346. doi: 10.3102/0002831211419649
- 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. (2011). The negative effect of school-average ability on science self-concept in the UK, the UK countries and the world: the big-fish-little-pond-effect for PISA 2006. *Educ. Psychol.* 31, 629–656. doi: 10.1080/01443410.2011.586416
- Pinheiro, J., Bates, D., Saikat, D., Sarkar, D., and R Core Team (2019). *nlme: Linear and Nonlinear Mixed Effects Models*. (Version R Package Version 3.1-142). Retrieved from: <https://CRAN.R-project.org/package=nlme>
- Pohlmann, B., and Möller, J. (2006). Vergleichseffekte auf kognitive, affektive und motivationale Variablen [Comparison effects on cognitive, affective, and motivational variables: an experimental approach]. *Z. Entwicklungspsychol. Pädagog. 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*. Retrieved from: <https://www.R-project.org/>
- Rheinberg, F. (1980). *Leistungsbewertung und Lernmotivation [Performance Evaluation and Learning Motivation]*. Göttingen: Hogrefe.
- Rheinberg, F., Vollmeyer, R., and Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen [QCM: a questionnaire to assess current motivation in learning situations]. *Diagnostica* 47, 57–66. doi: 10.1026//0012-1924.47.2.57
- Ross, M., and Wilson, A. E. (2003). Autobiographical memory and conceptions of self: getting better all the time. *Curr. Dir. Psychol. Sci.* 12, 66–69. doi: 10.1111/1467-8721.01228
- Ruble, D. N., Boggiano, A. K., Feldman, N. S., and Loeb, J. H. (1980). Developmental analysis of the role of social comparison in self-evaluation. *Dev. Psychol.* 16, 105–115. doi: 10.1037/0012-1649.16.2.105
- 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 levels of text comprehension," In *The Role of Interest in Learning and Development*, eds K. A. Renninger, S. Hidi, and A. Krapp (Hillsdale, NJ: 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
- 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, NJ: Lawrence Erlbaum Associates), 217–245.
- Seaton, M., Marsh, H. W., and Craven, R. G. (2009). Earning its place as a pan-human theory: universality of the big-fish-little-pond effect across 41 culturally and economically diverse countries. *J. Educ. Psychol.* 101, 403–419. doi: 10.1037/a0013838
- Shavelson, R. J., Hubner, J. J., and Stanton, G. C. (1976). Self-concept: validation of construct interpretations. *Rev. Educ. Res.* 46, 407. doi: 10.3102/00346543046003407
- Springer, L., Stanne, M. E., and Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. *Rev. Educ. Res.* 69, 21–51. doi: 10.3102/00346543069001021
- 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
- Valentine, J. C., DuBois, D. L., and Cooper, H. (2004). The relation between self-beliefs and academic achievement: a meta-analytic review. *Educ. Psychol.* 39, 111–133. doi: 10.1207/s15326985ep3902\_3
- Wang, Z. (2015). Examining big-fish-little-pond-effects across 49 countries: a multilevel latent variable modelling approach. *Educ. Psychol.* 35, 228–251. doi: 10.1080/01443410.2013.827155
- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. R., Arbreton, A. J. A., Freedman-Doan, C., et al. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: a 3-year study. *J. Educ. Psychol.* 89, 451–469. doi: 10.1037/0022-0663.89.3.451
- Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychol. Bull.* 90, 245–271. doi: 10.1037/0033-2909.90.2.245
- Wouters, S., De Fraine, B., Colpin, H., Van Damme, J., and Verschueren, K. (2012). The effect of track changes on the development of academic self-concept in high school: a dynamic test of the big-fish-little-pond effect. *J. Educ. Psychol.* 104, 793–805. doi: 10.1037/a0027732
- Zeidner, M., and Schleyer, E. J. (1999). The big-fish-little-pond effect for academic self-concept, test anxiety, and school grades in gifted children. *Contemp. Educ. Psychol.* 24, 305–329. doi: 10.1006/ceps.1998.0985

**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.

Copyright © 2020 Bosch and Wilbert. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.