

Contrast and Assimilation Effects on Task Interest in an Academic Learning Task

Jannis Bosch & Jürgen Wilbert

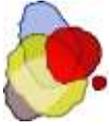
University of Potsdam, Germany

Article received 14 February / revised 18 May / accepted 2 August / available online 7 August

Abstract

Information on social comparison is one of the major factors used to evaluate academic achievement. The presence of Big-Fish-Little-Pond (BFLP) and Basking-In-Reflected-Glory (BIRG) effects of academic achievement on the self-concept have been extensively researched in various observational studies. Recent research suggests that these effects can also be transferred to motivational variables such as task interest. This paper uses an experimental paradigm to take a closer look at the mechanisms expected to be behind BFLP and BIRG effects: contrast and assimilation effects of task performance. The analyses are based on $N = 129$ primary education students who completed a computer-based learning task. During this task, participants received social comparative feedback that was experimentally manipulated based on 2x2 conditions: social position (high vs. low) and peer performance (high vs. low). Task interest was measured both before the start of the learning task and after it was finished. Results indicate a positive influence of high social position and high peer performance on the development of task interest from pre- to post-test compared to the low social position and peer performance conditions. Further, the hypothesized positive association between self-concept, task interest and performance in the academic learning task could also be shown.

Keywords: social comparison; feedback; interest; contrast effect; assimilation effect

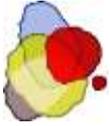


1. Theory

A lot of time during childhood, adolescence, and early adulthood is spent in institutionalized learning situations, such as school, vocational training, or university. Usually, these institutions are group-based and therefore social in nature. While these settings enable the use of group-based learning activities that can have several positive effects on learning achievement and attitudes towards learning (Springer, Stanne, & Donovan, 1999), they also facilitate the use of social comparison information in the evaluation of students' learning progress or achievement. Hence, information on social comparison is one of the major factors used to evaluate academic achievement in many educational contexts, both formally (e.g., grades) and informally (e.g., contact with peers). While social comparison can have desirable effects, it can also have considerable downsides, such as negative affect and lack of self-confidence, especially for lower-achieving students: “In classrooms characterized by frequent grades and public evaluation, students become focused on their ability and the distribution of ability in the classroom group. Many students not only come to believe that they lack ability but this perception also becomes shared among peers” (Ames, 1992, p. 264). This becomes even more relevant considering the increasing heterogeneity of academic learning groups that can be observed lately, due to the implementation of the “Convention of the Rights of Persons with Disabilities” in 2008 and the corresponding move to an inclusive school system in many countries. Processes of social comparison not only present a risk for low-achieving students to develop a negative self-concept, but can also influence academic achievement itself. Studies have shown that motivational (i.e. interest; Möller, Retelsdorf, Köller, & Marsh, 2011; Schiefele, Krapp, & Winteler, 1992) and self-related emotional (i.e. self-concept; Guay, Marsh, & Boivin, 2003; Huang, 2011; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Möller et al., 2011; Valentine, DuBois, & Cooper, 2004) variables play a very important role in the academic learning process. Further, the relation between academic achievement and both interest and self-concept seems to be bidirectional, as research reveals that low academic achievement in turn leads to a negative academic self-concept and hampers the development of interest in an academic topic (Denissen, Zarrett, & Eccles, 2007; Marsh et al., 2005), which can potentially lead to a vicious circle of negative self-affect, diminished interest and negative performance feedback in academic subjects. Hence, the establishment of motivating and emotionally stabilizing evaluation practices should also be a major goal to ensure successful institutionalized learning for every student.

To be able to do that, we need to analyse how the use of social comparison data in academic evaluation variously influences the interest and self-concept of students with heterogeneous prerequisites. Although there is ample evidence of the impact socially comparative feedback has on the development of an individual's self-concept, the role of interest and motivation has been considerably less well scrutinized. That said, theoretical models, as well as empirical evidence, propose a tight connection between self-concept and interest in an academic context (e.g. Fryer, 2015). Both constructs are domain-specific, and the associations between them always correspond to a specific domain (i.e. math self-concept is associated with math interest). Obviously, self-concept and interest are two distinctive constructs nonetheless: Interest, on one hand, is the motivational orientation of a person toward a certain object, activity, or area of knowledge that is independent of any direct external rewards and can therefore be described as domain-specific intrinsic motivation (Schiefele, 1992). Self-concept, on the other hand, consists of a person's perceptions of him- or herself. These perceptions are formed based on self-evaluation as well as the evaluation of significant others. That means that self-concept is self-ascribed competence in a specific domain (Marsh & Shavelson, 1985; Shavelson, Hubner, & Stanton, 1976). The Self-Determination Theory (Deci & Ryan, 1985), one of the most popular motivational theories, has identified feelings of competence in a specific domain as crucial for the establishment of intrinsic motivation in that same domain, suggesting a possible explanation for empirical connections between self-concept and interest found in several studies (e.g. Fryer, 2015; Schurtz, Pfof, Nagengast, & Artelt, 2014).

Taking the feeling of competence into focus as a bridge between self-concept and interest, empirical evidence for the effects of social comparison on self-concept will be presented in the following section. According to the Internal/External Frame of Reference Model by Marsh (I/E-Model; 1986) development of the self-concept is based on comparison of perceived competence in a specific domain with both internal (i.e. perceived competence in other domains) and external (i.e. perceived competence of other persons in that same

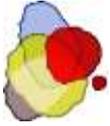


domain) frames of reference. While the former can be expected to have an overall neutral effect on the sum of self-concepts in different domains (e.g. a person perceives himself as good in math but bad in arts), the latter can have predominantly positive or negative effects on the self-concept in various domains. Hence, social comparison processes are a major factor for the development of a positive self-concept, especially when the evaluative atmosphere of the typical classroom makes it hard for students not to compare themselves with each other (Dijkstra, Kuyper, van der Werf, Buunk, & van der Zee, 2008). As stated earlier, social comparison processes can be detrimental to the development of a positive self-concept in low-achieving students. The results of a study by Dickhäuser and Galfe (2004) suggest that the direction of comparison (i.e. comparison with a better or worse performing peer) has an influence on the development of the self-concept. In their study, students were asked to whom they compared themselves after their math test scores were announced. Additionally, math self-concept was assessed before and after the math test. The results showed a negative effect of upward social comparison (i.e. comparison with a better performing peer) on math self-concept, while downward social comparison (i.e. comparison with a worse performing peer) tended to produce more positive effects. Hence, the predominance of upward comparison targets might be one factor responsible for the previously stated negative effects of social comparison on self-concept development in low-achieving students. This notion is further supported by two studies that investigated the effects of ability grouping on self-concept: The first study showed that ability grouping leads to decreased academic self-concept for those in higher ability groups compared to those in lower ability groups (Mulkey, Catsambis, Steelman, & Crain, 2005), while the second study found a similar effect of lowered academic self-concept for academically handicapped special school students who were integrated into mainstream schooling (Strang, Smith, & Rogers, 1978).

Therefore, the influence of social comparison on self-evaluation, and its effects on self-concept, cannot be properly understood without considering the composition of the reference group. The intimate nature of the process connecting self-evaluation and the attributes of social comparison groups have been outlined in the Inclusion/Exclusion Model (IEM; Bless & Schwarz, 2010; Schwarz & Bless, 1992). According to the IEM, evaluation of a particular stimulus is based on the stimulus itself and a standard of comparison. Because humans can never consider all potentially relevant information, both the stimulus itself and the standard of comparison are defined by the subset of information that is most accessible at the time of judgment. The model assumes two possible effects of the standard of comparison on the evaluation of the target: a *contrast* effect, i.e. the use of the standard of comparison as a reference against which to judge any characteristic of the target, and an *assimilation* effect, i.e. the transfer of characteristics of the standard of comparison to the target. The IEM covers both situations in which the target is superordinate to the standard of comparison, i.e. a group is judged in relation to an individual that is part of the group, and situations in which the target is subordinate to the standard of comparison, i.e. a member of the group is judged in relation to the group. If we apply the IEM to the process of self-evaluation, the target (e.g., a student) is part of a superordinate category (e.g., the class or school). While there are various sources of information that can be used for self-evaluation, information about other students in one's class (or school) is very salient and readily available in school contexts and will therefore most likely play a major role in school-based self-evaluations. Depending on which information is actually used to evaluate the self, both *contrast* effects, i.e. the usage of one's class (or school) as a reference against which to judge any of one's own characteristics, or *assimilation* effects, i.e. the transfer of characteristics of the students of one's class (or school) to oneself, can occur (see Bless & Schwarz, 2010 for a detailed discussion of influencing factors).

1.1 Social comparison effects in heterogeneous classrooms

The concept of *assimilation* and *contrast* effects supposedly lies at the core of one of the most popular phenomena in educational research. Based on their observation that students attending schools with a low average socio-economic status (SES) had higher academic self-concepts than equally performing students attending schools with a higher average SES, Marsh and Parker (1984) came up with the *big-fish-little-pond* (*BFLP*) and *basking-in-reflected-glory* (*BIRG*) effects. The *BFLP* effect is the transfer of a *contrast* effect to an institutionalized learning context; it states that self-evaluation of performance is based on the average



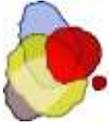
performance of other students in the same school, thus resulting in a higher academic self-concept for a student in a lower-performing school compared to an equally performing student in a higher-performing school. The *BIRG* effect is the transfer of an *assimilation* effect to an institutionalized learning context and has contrary expectations: A student in a high-performing school will have a higher academic self-concept than an equally performing student in a lower-performing school, because both students identify with their reference group and adapt their academic self-concept based on their group affiliation. However, because of a relatively stronger *BFLP* effect, the net result of the reference group's strength in terms of academic self-concept is still expected to be negative (Marsh, 1987). Since then, the *BFLP* effect on academic self-concept has been confirmed in school children from different countries (Marsh et al., 2005; Schurtz et al., 2014; Wouters, De Fraine, Colpin, Van Damme, & Verschueren, 2012), as well as in larger studies covering schools from various culturally and economically diverse countries (Chiu, 2012; Marsh & Hau, 2003; Seaton, Marsh, & Craven, 2009; Wang, 2015). Similar *BFLP* effects could also be found in young adults (Jonkmann, Becker, Marsh, Lüdtke, & Trautwein, 2012). Further, several studies have also shown empirical evidence for the existence of an - albeit usually smaller - *BIRG* effect in children and young adults (Huguet et al., 2009; Marsh, Kong, & Hau, 2000; Preckel & Brüll, 2010; Trautwein, Lüdtke, Marsh, & Nagy, 2009).

While the *BFLP* effect has mainly been investigated regarding academic self-concept, there are also several studies showing that the underlying processes may additionally affect other self-related constructs, as well as emotional variables. More specifically, these include *BFLP* effects on self-related constructs such as self-esteem (Jansen, Scherer, & Schroeders, 2015) and self-efficacy (Marsh, Trautwein, Lüdtke, & Köller, 2008), on general motivational variables such as educational and occupational aspirations (Marsh, 1991; Marsh & O'Mara, 2010; Nagengast & Marsh, 2011), on control expectations and strategies (Marsh et al., 2008), and on emotional variables such as test anxiety (Goetz, Preckel, Zeidner, & Schleyer, 2008; Zeidner & Schleyer, 1999) and emotionality (Goetz et al., 2008). Additionally, a similar *BFLP* effect could also be shown in regard to task interest (Köller, Schnabel, & Baumert, 2000; Schurtz et al., 2014; Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006), possibly reflecting the close connection between self-concept and interest. Trautwein and colleagues (2006) additionally tested for a *BIRG* effect on task interest. Although they did not find evidence for a *BIRG* effect on task interest, the failure to reproduce the original *BIRG* effect on self-concept renders those results somewhat fragile. Even though the presented results provide indirect evidence for the existence of *contrast* and *assimilation* effects in institutionalized education contexts, the non-experimental nature of the described studies still leaves some questions in relation to the assumed modes of action unanswered.

In order to be able to investigate the direct influence of social comparative performance feedback (i.e. *contrast* and *assimilation* effects) on task interest, experimental designs explicitly manipulating the performance feedback are necessary. To the author's knowledge, only one study used such an experimental design to show that *contrast* effects of performance feedback have a direct influence on task interest (Pohlmann & Möller, 2006). A previous experimental study by Wilbert, Grosche, and Gerdes (2010) could also show that social comparison feedback tends to increase task-specific motivation in high-achieving students while undermining low-achieving students' motivation. However, they did not control for the fact that high-achieving students usually receive positive feedback, while low-achieving students receive negative feedback.

1.2 Research question and hypotheses

In this paper, we want to further investigate the assumed modes of action behind the *BFLP* and *BIRG* effects: the direct influence *contrast* and *assimilation* effects in social comparative performance feedback have on students' interest in an academic learning task. The only study directly investigating the influence of social comparative performance feedback on task interest could find evidence for a *contrast* effect (Pohlmann & Möller, 2006). They did not, however, test for potential *assimilation* effects. Therefore, we want to investigate the direct effects of social comparative feedback on task interest to shed more light on the modes of action expected to be behind the popular *BFLP* and *BIRG* effects.



Hence, the research question is as follows: How do *contrast* and *assimilation* effects of social comparative performance feedback directly influence task interest in an academic learning task?

Based on the empirical results and theoretical considerations presented above, we expect both *contrast* and *assimilation* effects to influence interest in an academic learning task. That is, we expect a higher social position within a reference group (compared to a lower social position), as well as a higher peer performance (compared to lower peer performance) to be associated with higher interest in an academic learning task.

In order to achieve a clear differentiation between the quality of feedback and actual performance, and therefore to increase the internal validity, feedback related to both social position and peer performance was experimentally manipulated in this study. That means the performance feedback every participant received was independent of actual performance. To directly investigate potential *contrast* effects the percentile rank of performance feedback given was experimentally manipulated (high and low social position conditions). This was done to simulate the respective performance feedback high and low performing learners usually receive. Further, potential *assimilation* effects were investigated by experimentally manipulating the percentage of points the reference group scored (high and low peer performance conditions).

Additionally, we want to replicate previously presented positive associations between self-concept, interest and achievement in an academic learning task. Finally, the potential influences of *contrast* and *assimilation* effects on task enjoyment, a further variable strongly associated with task-interest, are explored.

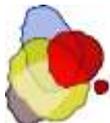
2. Method

2.1 Participants

Participants were 129 first year pedagogy students at University of Potsdam studying to become elementary school teachers. Seven participants were excluded from further analyses due to missing data or because they expressed doubt about the authenticity of the performance feedback, leaving us with 122 participants. Each participant was randomly assigned to one of 2x2 experimental conditions: 31 participants were in the high *social position* / high *peer performance* (SP+/PP+) condition, 30 participants in the low *social position* / high *peer performance* (SP-/PP+) condition, 32 participants were in the high *social position* / low *peer performance* (SP+/PP-) condition, and 29 participants were in the low *social position* / low *peer performance* (SP-/PP-) condition. 110 participants were female, 11 were male, and 1 did not report a gender. The age of the participants ranged from 18 to 37 ($M = 22.19$, $SD = 4.29$). All participants attended a lecture on inclusive education and were recruited during the class.

2.2 Learning task

The learning task we used in this study was presented to the participants as the “flag game.” Participants were told that the game consisted of two phases: a learning and a performance phase. During the learning phase, participants were presented with a map of the entire African continent and the country outlines for every country. For each of the 30 learning trials, the outlines of one country were highlighted, accompanied by the corresponding national flag. Each combination was presented for up to 10 seconds. Participants could shorten the presentation of each pair by pressing the space key. After completion of the learning phase, the performance phase was started. During the performance phase, the same 30 items were presented. This time, however, five different country outlines were highlighted with every flag, that is, four distractors in addition to the correct country. Participants were instructed to choose the correct country as quickly as possible by clicking on the corresponding country name. There was no time limit in the performance phase. After the first run, feedback was given depending on the respective feedback condition. Then, a second run comprising of



another learning and performance phase with the same 30 items was started. After the second run, no feedback was given.

To avoid sequencing effects between conditions, trial sequences were pseudo-randomized separately for each phase. Distractor items were taken from the item pool, balanced and pseudo-randomized separately for each performance phase. Hence, every participant was presented with the exact same trial sequences and distractor items.

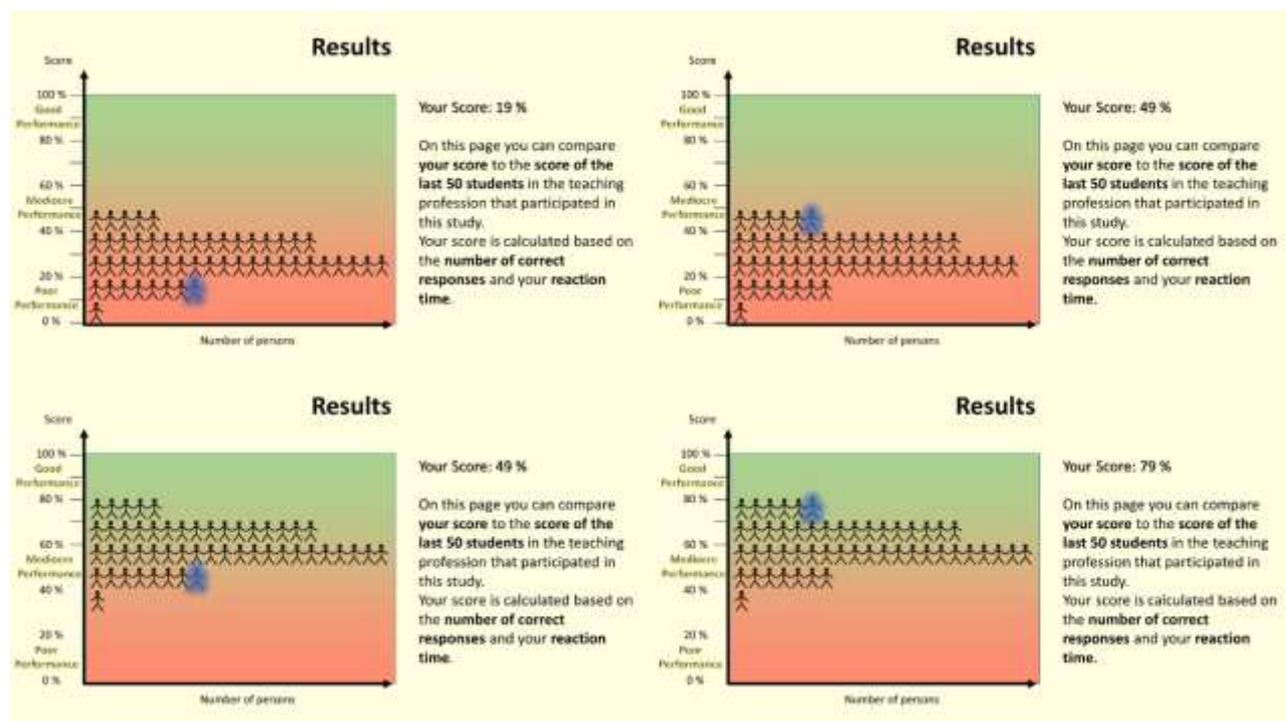
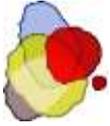


Figure 1. Feedback slides for all four experimental conditions (feedback slides were originally in German and were translated for this article). The slides on the left were presented to the respective low social position conditions (top side: low social position/low peer performance, bottom side: low social position/high peer performance), while the slides on the right were presented to the respective high social position conditions (top side: high social position/low peer performance, bottom side: high social position/high peer performance).

2.3 Experimental feedback conditions

During instruction, participants were told that they would receive information about their performance compared to the last 50 students of the same degree course. However, feedback did not reflect their real performance, but was given based on the respective feedback condition. To test for *contrast* and *assimilation* effects on task interest, we manipulated both *social position* (SP; high vs. low) and *peer performance* (PP; high vs. low) in a 2x2 design. Participants were told their performance rating could range from 0 to 100%, based on both correctness and the speed of their responses. Response speed was included to make it harder for the participants to predict their own performance. Performance feedback was shown graphically, with 50 stick figures in a graph representing the scores of the reference group and a highlighted stick figure representing the participants' own score (see Figure 1). Scores for the reference group were approximately normally distributed and ranged between 0 and 50% in the low *peer performance* conditions. In the high *peer performance* conditions, the reference group score was simply raised by 30% for every reference group member, resulting in values between 30% and 80%. Based on the respective *social position* condition, the participant's individual



score was placed either at the top or the bottom of the distribution. This resulted in individual scores of 79% (high *social position* / high *peer performance*), 49% (high *social position* / low *peer performance* and low *social position* / high *peer performance*), or 19% (low *social position* / low *peer performance*).

2.4 Measures

2.4.1 Self-concept

The self-concept scale consists of four items: two positive (e.g., “Usually I have no trouble with learning tasks”) and two negative (e.g., “Dealing with learning tasks isn’t one of my strengths”) statements about the participant’s self-rated ability to resolve learning tasks. Participants rated the items on a Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree). Negative items were reversed and a mean self-concept score was calculated. Cronbach’s alpha of the scale was $\alpha = .84$.

2.4.2 Task interest (pre- and post-test)

The interest scale consists of eight items: four items concerning interest in the task itself (e.g., “I like learning tasks such as this one”) and four items concerning interest in the content of the task (e.g., “I’m interested in country flags”). The items were adapted versions of an interest scale from a German motivation questionnaire (Rheinberg, Vollmeyer, & Burns, 2001). Each item was rated on a Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree). Pre-test interest had a Cronbach’s alpha of .87, while post-test interest had a Cronbach’s alpha of $\alpha = .90$.

2.4.3 Task enjoyment

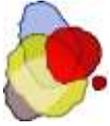
Task enjoyment was measured by means of a single item at the end of the study (“Did you enjoy the study?”). The item was rated on a Likert-type scale from 1 (definitely no) to 4 (definitely yes).

2.4.4 Task performance (pre- and post-test)

Pre-test performance was the proportion of correctly answered multiple-choice items in the first performance phase, while post-test performance was the percentage of correctly answered items in the second performance phase. Both performance phases consisted of the same 30 items.

2.5 Procedure

Participants were tested in groups of four to six persons in our learning laboratory. They were welcomed and informed about the upcoming procedure. Written informed consent was obtained from every participant prior to testing. Each participant was placed in front of a separate computer. Computers were properly shielded to avoid participants seeing each other’s screens. Firstly, the self-concept questionnaire was administered. During the computer-based instruction, participants were informed about the upcoming learning task and how feedback about their performance would be provided. After completion of the instruction, a questionnaire concerning task interest was administered. When all participants finished filling out the questionnaire, the computer-based learning task was begun. Afterwards, the task interest questionnaire was administered a second time to investigate any potential changes. In the end, participants answered two questions concerning whether they were motivated to participate in similar studies and an open question that functioned as a manipulation check. The true nature of the study and the experimental manipulation was then explained to all participants.



2.6 Statistical analyses

To test the previously presented hypotheses that there are *contrast* and *assimilation* effects on domain-specific interest, we followed the suggestion by Everitt and Hothorn (2011, p. 232) and used linear mixed models with fixed and random effects. This was done because regular linear regression is unable to deal with data from repeated measures. Maximum likelihood (ML) estimation was used for all models. Based on the suggestions by Nakagawa and Schielzeth (2013), marginal and conditional R^2 were reported in order to determine variance explained by fixed factors and by both fixed and random factors, respectively. Model 1 contained the subject as a random intercept factor and pre- and post-test interest as Level 1 variables. Contrasts for the *measurement time* factor were coded 0 for the pre- and 1 for the post-test measurements. For Model 2, we added the Level 2 variable *social position* and the interaction with *measurement time*. Then, *peer performance*, as well as interactions with *measurement time* and *social position* were added for Model 3. Finally, we also added the standardized ($M = 0$, $SD = 1$) self-concept and interaction terms to the other predictors in order to investigate possible moderating effects in Model 4. The contrasts were effect-coded as -1 for low *social position* or *peer performance* and +1 for high *social position* or *peer performance*, respectively. To determine the model best fitting our data, likelihood ratio (LR) tests comparing each model to the previous one were calculated. This was done to test whether the subsequent addition of predictors significantly increased variance explained.

To investigate the possible effects of experimental conditions on learning task performance, corresponding models were also calculated with task performance as the dependant variable.

Because task enjoyment was only retrieved at a single point in time at the end of the study, generalized least squares (GLS) models with ML estimation and only one measurement per subject were calculated to investigate *BFLP* effects on task enjoyment. In order to determine explained variance in the GLS models, Likelihood Ratio R^2 was calculated as suggested by Magee (1990). This time, an intercept-only model with task enjoyment as the dependent variable was used as the base Model 1. Again, *social position* was included in Model 2, while *peer performance* and the interaction between *social position* and *peer performance* were added to Model 3 and standardized self-concept and corresponding interactions to Model 4. Once again, LR tests were calculated to compare the four models.

All calculations were carried out using the statistical software *R* with *nlme* and *MuMin* packages.

3. Results

Table 1 shows intercorrelations and descriptive statistics for all observed variables. As can be seen, self-concept is positively associated with pre- and post-test interest. Associations between self-concept and pre- and post-test performance was also positive, but only significant in the latter case. Further, pre-test interest showed a significant positive association to pre- and post-test performance. Comparable associations could be found between post-test interest and pre- and post-test performance.

Hence, results support the hypothesized associations between interest, self-concept and performance.



Table 1

Descriptive statistics and intercorrelations for all observed variables

	Self- Concept	Interest Pre-Test	Interest Post-Test	Interest Diff.	Perf. Pre-Test	Perf. Post-Test	Perf. Diff.	Enjoy
Self-Concept	1							
Interest Pre-Test	.52**	1						
Interest Post-Test	.39**	.87**	1					
Interest Diff.	-.16	-.07	.44**	1				
Perf. Pre-Test	.13	.23*	.31**	.20*	1			
Perf. Post-Test	.23*	.20*	.25*	.14	.74**	1		
Perf. Diff.	.15	-.02	-.06	-.08	-.30**	.42**	1	
Enjoy	.09	.24*	.29**	.15	.01	.04	.04	1
<i>M</i>	4.39	3.81	3.88	0.07	.51	.71	.20	3.40
<i>SD</i>	1.03	1.16	1.28	0.64	1.15	1.16	1.11	0.65
<i>N</i>	122	122	122	122	122	122	122	121

Note. ** $p < .01$, * $p < .05$; Interest Diff. = difference between interest pre- and post-test, Perf. Pre-Test = performance pre-test, Perf. Post-Test = performance post-test, Perf. Diff. = difference between performance pre- and post-test, Enjoy = enjoyment

Table 2

Descriptive statistics for all observed variables for each experimental condition

	SP + PP +	SP – PP +	SP + PP –	SP – PP –
Self-Concept				
Pre-Test	4.47 (1.04)	4.21 (0.95)	4.31 (1.16)	4.59 (0.96)
Interest				
Pre-Test	4.02 (1.14)	3.44 (1.13)	3.91 (1.11)	3.85 (1.23)
Post-Test	4.39 (1.12)	3.43 (1.33)	3.94 (1.17)	3.73 (1.39)
Difference	0.37 (0.64)	- 0.01 (0.69)	0.03 (0.57)	- 0.12 (0.58)
Performance				
Pre-Test	0.53 (0.16)	0.49 (0.15)	0.52 (0.16)	0.50 (0.12)
Post-Test	0.75 (0.16)	0.71 (0.16)	0.70 (0.17)	0.69 (0.12)
Difference	0.21 (0.12)	0.22 (0.12)	0.18 (0.11)	0.19 (0.10)
Enjoyment				
Post-Test	3.61 (0.50)	3.24 (0.69)	3.44 (0.72)	3.28 (0.65)

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

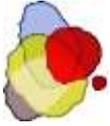


Table 2 shows the descriptive results separately for each experimental condition. One-factorial ANOVAs with the experimental condition as independent variable showed no significant differences between groups for self-concept ($F_{3, 118} = 0.78, p = .51$), pre-test interest ($F_{3, 118} = 1.45, p = .23$) or pre-test performance ($F_{3, 118} = 0.53, p = .67$). Hence, there were no significant pre-test differences between the groups in relevant study variables.

3.1 Contrast and assimilation effects on task interest

Table 3 shows the model fit and parameter estimates for all models predicting task interest. Model 2 including *measurement time*, *social position* and the interaction between *social position* and *measurement time* showed significantly better model fit compared to Model 1 only containing *measurement time* ($LR = 9.79, p < .01$). Inclusion of *peer performance* and interactions (Model 3) did increase the model fit, albeit not significantly ($LR = 7.72, p = .10$). Addition of the self-concept and interactions to all previous predictors (Model 4) further increased the model fit ($LR = 38.29, p < .001$).

As can be seen in Model 1, *measurement time* did not significantly predict task interest, showing that there were no differences in task interest between *measurement times* across all experimental conditions.

Model 2 contained the hypothesized interaction between *social position* and *measurement time*, reflecting a more positive development of task interest in the high *social position* conditions, compared to the low *social position* conditions.

Model 3 further contained the hypothesized interaction between *measurement time* and *peer performance* as a significant predictor, once again reflecting a more positive interest development in the high *peer performance* compared to the low *peer performance* conditions.

Model 4 additionally included *self-concept* as a significant predictor, showing a strong association between *self-concept* and task interest. The interaction between *self-concept* and *measurement time*, however, was small and did not reach significance. Self-concept also did not moderate interactions between *social position* and *measurement time* or between *peer performance* and *measurement time*.

Hence, even though task interest did not change from pre- to post-test measure across the whole sample, there were interactions between *social position* and *measurement time* as well as between *peer position* and *measurement time*, reflecting a more positive development of task interest in the high *social position* and high *peer performance* conditions compared to the low *social position* and low *peer performance* conditions, respectively. Additionally, *self-concept* was associated with pre-test task interest, but did not influence the effect of experimental conditions on the development of task interest. Therefore, our results support the hypothesized influence of *contrast* and *assimilation* effects on task interest.

3.2 Contrast and assimilation effects on task performance

Table 4 shows the model fit and parameter estimates for all models predicting task performance. Neither Model 2 containing *social position* and the interaction between *social position* and *measurement time* ($LR = 1.29, p = .52$), nor Model 3 additionally containing *peer performance* and corresponding interactions ($LR = 2.91, p = .57$), nor Model 4 further containing self-concept and corresponding interactions ($LR = 10.13, p = .26$) showed significantly better model fit than the model containing only repeated measurements.

Model 1 did contain a significant predictor for *measurement time*, showing that our participants performed significantly better in the second compared to the first run. These differences in task performance were influenced by neither *social position* nor *peer performance*.



Table 3

Fixed effects for mixed models predicting interest ($N = 122$)

Parameter	Unstandardized Estimate b (SE)			
	Model 1	Model 2	Model 3	Model 4
Intercept	3.81** (0.111)	3.80** (0.109)	3.81** (0.109)	3.81** (0.099)
Level 1				
MT	0.07 (0.058)	0.07 (0.057)	0.07 (0.056)	0.07 (0.057)
Level 2				
SP		0.16 (0.109)	0.16 (0.109)	0.14 (0.099)
PP			- 0.08 (0.109)	- 0.05 (0.099)
SC				0.58** (0.098)
Interactions				
MT x SP		0.13* (0.057)	0.13* (0.056)	0.13* (0.057)
MT x PP			0.12* (0.056)	0.12* (0.057)
MT x SC				- 0.09 (0.056)
SP x PP			0.13 (0.109)	0.05 (0.099)
SP x SC				0.02 (0.098)
PP x SC				0.14 (0.098)
MT x SP x PP			0.06 (0.056)	0.07 (0.057)
MT x SP x SC				- 0.04 (0.056)
MT x PP x SC				0.06 (0.056)
SP x PP x SC				0.01 (0.098)
MT x SP x PP x SC				- 0.01 (0.056)
Model Indices				
<i>logLik</i>	- 311.13	- 306.24	- 302.38	- 283.24
<i>df</i>	4	6	10	18
<i>R²_m</i>	.00	.04	.06	.27
<i>R²_c</i>	.86	.87	.88	.88

Note. ** $p < .01$, * $p < .05$; MT = measurement time (0/1), SP = social position (-1/+1), PP = peer performance (-1/+1), SC = self-concept (centered), logLik = logarithmized Likelihood of the model, df = degrees of freedom, R^2_m = marginal R^2 , R^2_c = conditional R^2

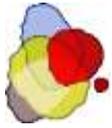


Table 4

Fixed effects for mixed models predicting task performance ($N = 122$)

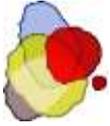
Parameter	Unstandardized Estimate b (SE)			
	Model 1	Model 2	Model 3	Model 4
Intercept	0.51** (0.013)	0.51** (0.014)	0.51** (0.014)	0.61** (0.014)
Level 1				
MT	0.20** (0.010)	0.20** (0.010)	0.20** (0.010)	0.10** (0.010)
Level 2				
SP		0.02 (0.014)	0.02 (0.014)	0.02 (0.014)
PP			0.01 (0.014)	0.00 (0.014)
SC				0.02 (0.014)
Interactions				
MT x SP		0.00 (0.010)	0.00 (0.010)	0.00 (0.010)
MT x PP			0.01 (0.010)	0.01 (0.010)
MT x SC				0.02 (0.010)
SP x PP			0.01 (0.014)	0.00 (0.014)
SP x SC				0.02 (0.014)
PP x SC				-0.01 (0.014)
MT x SP x PP			0.00 (0.010)	0.00 (0.010)
MT x SP x SC				0.00 (0.010)
MT x PP x SC				0.00 (0.005)
SP x PP x SC				0.01 (0.014)
MT x SP x PP x SC				0.00 (0.010)
Model Indices				
<i>logLik</i>	161.70	162.34	163.80	168.86
<i>df</i>	4	6	10	18
<i>R²_m</i>	.31	.31	.32	.35
<i>R²_c</i>	.82	.82	.82	.82

Note. ** $p < .01$, * $p < .05$; MT = measurement time (0/1), SP = social position (-1/+1), PP = peer performance (-1/+1), SC = self-concept (centered), logLik = logarithmized Likelihood of the model, df = degrees of freedom, R^2_m = marginal R^2 , R^2_c = conditional R^2

3.3 Contrast and assimilation effects on task enjoyment

Table 5 shows the model fit and parameter estimates for all four models predicting task enjoyment. Model 2 containing the *social position* showed significantly better model fit than the intercept-only Model 1 ($LR = 5.15$, $p < .05$), while inclusion of *peer performance* and the interaction between *social position* and *peer performance* in Model 3 did not further increase the model fit ($LR = 1.25$, $p = .54$). Inclusion of self-concept (Model 4) also did not significantly increase model fit ($LR = 2.63$, $p = .62$).

Model 2 included *social position* as a significant predictor, while in Model 3 neither *peer performance* nor the interaction between *social position* and *peer performance* significantly predicted task enjoyment. In Model 4, self-concept was also not a significant predictor.



Hence, while the *social position* condition did influence task enjoyment in the hypothesized direction (i.e. feedback of a high *social position* led to higher task enjoyment compared to feedback of a low *social position*), there was no effect of the *peer performance* condition on task enjoyment.

Table 5

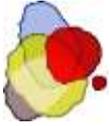
Fixed effects for mixed models predicting task enjoyment (N = 121)

Parameter	Unstandardized Estimate b (SE)			
	Model 1	Model 2	Model 3	Model 4
Intercept	3.40** (0.059)	3.39** (0.058)	3.39** (0.059)	3.40** (0.060)
Level 2				
SP		0.13* (0.058)	0.13* (0.059)	0.12* (0.060)
PP			0.04 (0.059)	0.04 (0.060)
SC				0.05 (0.059)
SP x PP			0.05 (0.059)	0.04 (0.060)
SP x SC				- 0.03 (0.059)
PP x SC				0.04 (0.059)
SP x PP x SC				- 0.06 (0.059)
Model Indices				
<i>logLik</i>	- 119.37	- 116.80	- 116.17	- 114.86
<i>df</i>	3	4	6	10
<i>R²_{LR}</i>	.00	.04	.05	.07

Note. ** $p < .01$, * $p < .05$; SP = social position (-1/+1), PP = peer performance (-1/+1), SC = self-concept (centered), *logLik* = logarithmized Likelihood of the model, *df* = degrees of freedom, *R²_{LR}* = Likelihood Ratio *R²*

4. Discussion

In this study, we tested whether *contrast* and *assimilation* effects on task interest, which are thought to be the basis of popular *BFLP* and *BIRG* effects, could be replicated in an experimental paradigm. As expected, the feedback of both a high *social position* and high *peer performance* had a positive influence on the change in our participants' interest over the course of this study (compared to low *social position* and *peer performance*, respectively). When considering the sample as a whole, there was only a minor, non-significant increase in task interest between pre- and post-test. However, when taking a closer look at group differences in interest development, a pattern matching the hypothesized additive effects of *contrast* and *assimilation* effects could be found. While feedback of high *social position* / high *peer performance* showed a clear increase, task interest remained virtually unchanged in both intermediate groups (i.e. feedback of high *social position* / low *peer performance* and feedback of low *social position* / high *peer performance*). Further, feedback of low *social position* / low *peer performance* even caused a decrease in task interest. Hence, this study provides additional experimental evidence for the hypothesis that *contrast* and *assimilation* effects of performance (as described in the IEM by Schwarz and Bless, 1992), and subsequent positive or negative performance feedback are likely to be at the core of *BFLP* and *BIRG* effects in task interest.



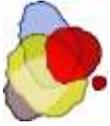
However, when it comes to task interest social comparative feedback is a zero-sum game. Not only are there winners, but also a rather consistent group whose interest is systematically undermined by continuing processes of negative social comparative evaluation, which could lead to further decline of academic achievement (Möller et al., 2011; Schiefele, 1992). This tends to be less of a problem in relatively homogeneous groups, where shortcomings in prior knowledge can more easily be overcome by increased effort. In rather heterogeneous groups, however, the influence of effort on academic evaluation is much lower, resulting in continuing negative social-comparative evaluation and therefore diminishing academic interest in slower learners. Hence, the focus on grades as the most popular form of formal feedback in institutionalized learning settings should be re-evaluated in order to facilitate effective learning for all students. Butler (1992), for example, could show that task interest for sixth graders differed depending on whether the instruction focused on task mastery or social comparison as well as on their self-rated performance. Subjects who rated their performance highly showed higher interest in the social comparison condition, while subjects with low self-rated performance showed higher interest in the mastery condition. The results from Harks, Rakoczy, Hattie, Besser, and Klieme (2014) point in a similar direction: Process-oriented feedback was perceived as more useful by ninth graders than grade-oriented feedback and also had a positive effect on the development of mathematics achievement and interest.

Feedback of a high *social position* did additionally lead to the hypothesized increase in task enjoyment compared to a low *social position*, suggesting a similar *contrast effect* of task performance on task enjoyment. *Peer performance* did not influence task enjoyment in the hypothesized direction, however, not supporting our hypothesis of an *assimilation* effect in task enjoyment. Hence, an effect of social comparative performance feedback on task enjoyment similar to the one on task interest could only be partially confirmed.

Further, we explored associations between self-concept, task interest and task performance and potential reference group effects on task interest. We were able to confirm the hypothesized positive associations between self-concept, interest and performance in an academic learning task. Hence, a positive self-concept usually coincided with higher interest as well as performance in our learning task.

4.1 Limitations

The experimental paradigm of this study was chosen to increase internal validity and further help to shed light on the exact mechanisms behind the *contrast* and *assimilation* effects of academic achievement. The downside is that the learning task in this study was strictly linear and lacked several dimensions of classroom learning (e.g. cooperation, teacher-student interaction, etc.). Therefore, experimental studies with learning situations more closely resembling those of classroom learning are necessary in order to gain a better understanding of the influence of *contrast* and *assimilation* effects on interest in actual school settings. Especially the operationalization of the two different *peer performance* conditions focuses on criterial differences in test score and lacks important aspects of the *BIRG* effect (e.g. being chosen into a highly valued group; highly visible selection process) that have been suggested (Marsh, 1987; Marsh, Kong, & Hau 2000). Further, pre-/post-differences in interest between the two conditions with individual scores of 49 % (i.e. high *social position* / low *peer performance* and low *social position* / high *peer performance*) were relatively small, possibly hinting at effects from the criterial score on interest that could not be controlled within the design of this study. Additionally, self-concept was only measured once at the beginning of the study. The addition of a post-learning-task measurement of self-concept would have enabled a deeper investigation of potential interactions between self-concept and task interest.

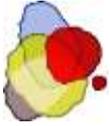


Keypoints

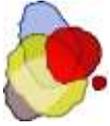
- Self-concept, interest and performance show hypothesized positive associations in an academic learning task
- Both high *social position* and high criterial *peer performance* could be shown to positively influence the development of task interest in an experimental paradigm
- Social comparative feedback seems to promote interest only for those receiving feedback above average, while having contrary effects for those receiving feedback below average

References

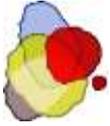
- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, *84*(3), 261–271. <https://doi.org/10.1037/0022-0663.84.3.261>
- Bless, H., & Schwarz, N. (2010). Mental Construal and the Emergence of Assimilation and Contrast Effects. *Advances in Experimental Social Psychology*, *42*, 319–373. [https://doi.org/10.1016/S0065-2601\(10\)42006-7](https://doi.org/10.1016/S0065-2601(10)42006-7)
- Butler, R. (1992). What young people want to know when: Effects of mastery and ability goals on interest in different kinds of social comparisons. *Journal of Personality and Social Psychology*, *62*(6), 934–943. <https://doi.org/10.1037/0022-3514.62.6.934>
- Chiu, M.-S. (2012). The internal/external frame of reference model, big-fish-little-pond effect, and combined model for mathematics and science. *Journal of Educational Psychology*, *104*(1), 87–107. <https://doi.org/10.1037/a0025734>
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.
- Denissen, J. J. A., Zarrett, N. R., & 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 Development*, *78*(2), 430–447. <https://doi.org/10.1111/j.1467-8624.2007.01007.x>
- Dickhäuser, O., & Galfe, E. (2004). Besser als ..., schlechter als... Leistungsbezogene Vergleichsprozesse in der Grundschule. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *36*(1).
- Dijkstra, P., Kuyper, H., van der Werf, G., Buunk, A. P., & van der Zee, Y. G. (2008). Social Comparison in the Classroom: A Review. *Review of Educational Research*, *78*(4), 828–879. <https://doi.org/10.3102/0034654308321210>
- Everitt, B., & Hothorn, T. (2011). *An Introduction to Applied Multivariate Analysis with R*. New York, NY: Springer New York. <https://doi.org/10.1007/978-1-4419-9650-3>
- Fryer, L. K. (2015). Predicting self-concept, interest and achievement for first-year students: The seeds of lifelong learning. *Learning and Individual Differences*, *38*, 107–114. <https://doi.org/10.1016/j.lindif.2015.01.007>
- Goetz, T., Preckel, F., Zeidner, M., & Schleyer, E. (2008). Big fish in big ponds: A multilevel analysis of test anxiety and achievement in special gifted classes. *Anxiety, Stress & Coping*, *21*(2), 185–198. <https://doi.org/10.1080/10615800701628827>
- Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, *95*(1), 124–136. <https://doi.org/10.1037/0022-0663.95.1.124>
- Harks, B., Rakoczy, K., Hattie, J., Besser, M., & Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: the role of feedback's perceived usefulness. *Educational Psychology*, *34*(3), 269–290. <https://doi.org/10.1080/01443410.2013.785384>
- Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology*, *49*(5), 505–528. <https://doi.org/10.1016/j.jsp.2011.07.001>



- Huguet, P., Dumas, F., Marsh, H., Régner, I., Wheeler, L., Suls, J., ... Nezlek, J. (2009). Clarifying the role of social comparison in the big-fish–little-pond effect (BFLPE): An integrative study. *Journal of Personality and Social Psychology*, *97*(1), 156–170. <https://doi.org/10.1037/a0015558>
- Jansen, M., Scherer, R., & Schroeders, U. (2015). Students' self-concept and self-efficacy in the sciences: Differential relations to antecedents and educational outcomes. *Contemporary Educational Psychology*, *41*, 13–24. <https://doi.org/10.1016/j.cedpsych.2014.11.002>
- Jonkmann, K., Becker, M., Marsh, H. W., Lüdtke, O., & Trautwein, U. (2012). Personality traits moderate the Big-Fish–Little-Pond Effect of academic self-concept. *Learning and Individual Differences*, *22*(6), 736–746. <https://doi.org/10.1016/j.lindif.2012.07.020>
- Köller, O., Schnabel, K. U., & Baumert, J. (2000). Der Einfluß der Leistungsstärke von Schulen auf das fachspezifische Selbstkonzept der Begabung und das Interesse. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *32*(2), 70–80. <https://doi.org/10.1026//0049-8637.32.2.70>
- Magee, L. (1990). R 2 Measures Based on Wald and Likelihood Ratio Joint Significance Tests. *The American Statistician*, *44*(3), 250. <https://doi.org/10.2307/2685352>
- Marsh, H. W. (1986). Verbal and Math Self-Concepts: An Internal/External Frame of Reference Model. *American Educational Research Journal*, *23*(1), 129149. <https://doi.org/10.3102/00028312023001129>
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, *79*(3), 280–295. <https://doi.org/10.1037/0022-0663.79.3.280>
- Marsh, H. W. (1991). Failure of High-Ability High Schools to Deliver Academic Benefits Commensurate with Their Students' Ability Levels. *American Educational Research Journal*, *28*(2), 445. <https://doi.org/10.2307/1162948>
- Marsh, H. W., & 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. *American Psychologist*, *58*(5), 364–376. <https://doi.org/10.1037/0003-066X.58.5.364>
- Marsh, H. W., Kong, C.-K., & 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. *Journal of Personality and Social Psychology*, *78*(2), 337–349. <https://doi.org/10.1037//0022-3514.78.2.337>
- Marsh, H. W., & O'Mara, A. J. (2010). Long-Term Total Negative Effects of School-Average Ability on Diverse Educational Outcomes: Direct and Indirect Effects of the Big-Fish-Little-Pond Effect. *Zeitschrift für Pädagogische Psychologie*, *24*(1), 51–72. <https://doi.org/10.1024/1010-0652/a000004>
- Marsh, H. W., & 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? *Journal of Personality and Social Psychology*, *47*(1), 213–231. <https://doi.org/10.1037/0022-3514.47.1.213>
- Marsh, H. W., & Shavelson, R. (1985). Self-Concept: Its Multifaceted, Hierarchical Structure. *Educational Psychologist*, *20*(3), 107–123. https://doi.org/10.1207/s15326985ep2003_1
- Marsh, H. W., Trautwein, U., Lüdtke, O., & Köller, O. (2008). Social comparison and big-fish-little-pond effects on self-concept and other self-belief constructs: Role of generalized and specific others. *Journal of Educational Psychology*, *100*(3), 510–524. <https://doi.org/10.1037/0022-0663.100.3.510>
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic Self-Concept, Interest, Grades, and Standardized Test Scores: Reciprocal Effects Models of Causal Ordering. *Child Development*, *76*(2), 397–416. <https://doi.org/10.1111/j.1467-8624.2005.00853.x>
- Möller, J., Retelsdorf, J., Köller, O., & Marsh, H. W. (2011). The Reciprocal Internal/External Frame of Reference Model: An Integration of Models of Relations Between Academic Achievement and Self-Concept. *American Educational Research Journal*, *48*(6), 1315–1346. <https://doi.org/10.3102/00028312111419649>
- Mulkey, L. M., Catsambis, S., Steelman, L. C., & Crain, R. L. (2005). The long-term effects of ability grouping in mathematics: A national investigation. *Social Psychology of Education*, *8*(2), 137–177. <https://doi.org/10.1007/s11218-005-4014-6>



- Nagengast, B., & 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. *Educational Psychology, 31*(5), 629–656. <https://doi.org/10.1080/01443410.2011.586416>
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>
- Pohlmann, B., & Möller, J. (2006). Vergleichseffekte auf kognitive, affektive und motivationale Variablen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 38*(2), 79–87. <https://doi.org/10.1026/0049-8637.38.2.79>
- Preckel, F., & Brüll, M. (2010). The benefit of being a big fish in a big pond: Contrast and assimilation effects on academic self-concept. *Learning and Individual Differences, 20*(5), 522–531. <https://doi.org/10.1016/j.lindif.2009.12.007>
- Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica, 47*(2), 57–66. <https://doi.org/10.1026/0012-1924.47.2.57>
- Schiefele, U. (1992). Topic Interest and Levels of Text Comprehension. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The Role of Interest in Learning and Development* (pp. 151–182). Hillsdale, N.J: L. Erlbaum Associates.
- Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a Predictor of Academic Achievement: A Meta-Analysis of Research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The Role of Interest in Learning and Development* (pp. 183–212). Hillsdale, N.J: L. Erlbaum Associates.
- Schurtz, I. M., Pfost, M., Nagengast, B., & Artelt, C. (2014). Impact of social and dimensional comparisons on student's mathematical and English subject-interest at the beginning of secondary school. *Learning and Instruction, 34*, 32–41. <https://doi.org/10.1016/j.learninstruc.2014.08.001>
- Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: Assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217–245). Hillsdale, N.J: Lawrence Erlbaum Associates.
- Seaton, M., Marsh, H. W., & 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. *Journal of Educational Psychology, 101*(2), 403–419. <https://doi.org/10.1037/a0013838>
- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-Concept: Validation of Construct Interpretations. *Review of Educational Research, 46*(3), 407. <https://doi.org/10.2307/1170010>
- Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of Small-Group Learning on Undergraduates in Science, Mathematics, Engineering, and Technology: A Meta-Analysis. *Review of Educational Research, 69*(1), 21–51. <https://doi.org/10.3102/00346543069001021>
- Strang, L., Smith, M. D., & Rogers, C. M. (1978). Social comparison, multiple reference groups, and the self-concepts of academically handicapped children before and after mainstreaming. *Journal of Educational Psychology, 70*(4), 487–497.
- Trautwein, U., Lüdtke, O., Marsh, H. W., Köller, O., & Baumert, J. (2006). Tracking, grading, and student motivation: Using group composition and status to predict self-concept and interest in ninth-grade mathematics. *Journal of Educational Psychology, 98*(4), 788–806. <https://doi.org/10.1037/0022-0663.98.4.788>
- Trautwein, U., Lüdtke, O., Marsh, H. W., & Nagy, G. (2009). Within-school social comparison: How students perceive the standing of their class predicts academic self-concept. *Journal of Educational Psychology, 101*(4), 853–866. <https://doi.org/10.1037/a0016306>
- Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The Relation Between Self-Beliefs and Academic Achievement: A Meta-Analytic Review. *Educational Psychologist, 39*(2), 111–133. https://doi.org/10.1207/s15326985ep3902_3
- Wang, Z. (2015). Examining big-fish-little-pond-effects across 49 countries: a multilevel latent variable modelling approach. *Educational Psychology, 35*(2), 228–251. <https://doi.org/10.1080/01443410.2013.827155>



- Wilbert, J., Grosche, M., & Gerdes, H. (2010). Effects of Evaluative Feedback on Rate of Learning and Task Motivation: An Analogue Experiment. *Learning and Disabilities: A Contemporary Journal*, 8(2), 43–52.
- Wouters, S., De Fraine, B., Colpin, H., Van Damme, J., & 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. *Journal of Educational Psychology*, 104(3), 793–805. <https://doi.org/10.1037/a0027732>
- Zeidner, M., & Schleyer, E. J. (1999). The Big-Fish–Little-Pond Effect for Academic Self-Concept, Test Anxiety, and School Grades in Gifted Children. *Contemporary Educational Psychology*, 24(4), 305–329. <https://doi.org/10.1006/ceps.1998.0985>