



Classroom effect on primary school students' self-concept in literacy and mathematics

Kati Vasalampi¹  · Eija Pakarinen²  · Minna Torppa²  · Jaana Viljaranta³  ·
Marja-Kristiina Lerkkanen²  · Anna-Maija Poikkeus² 

Received: 23 April 2018 / Revised: 6 August 2019 / Accepted: 23 September 2019 /

Published online: 26 November 2019

© The Author(s) 2019

Abstract

According to the big-fish-little-pond effect (BFLP) model, high individual academic performance in a particular subject is related to high self-concept in that subject, whereas high average classroom performance has a negative effect on self-concept. In the present study, data from Finnish primary school students in grade 3 (504 students), grade 4 (487 students), and grade 6 (365 students) are used to examine whether the assumptions of the BFLP effect model hold already in primary school. Furthermore, we examined gender differences in BFLP effect. The results showed that as expected students' high performance in literacy and in mathematics was related to high self-concept in the same subject. Support for the negative classroom effect was small and it depended on the school subject and student's gender. That is, a high average classroom performance already in grade 3 had a negative but small effect on boys' self-concept in mathematics. In literacy and among girls, only little support was found for the negative classroom effect.

Keywords The big-fish-little-pond effect · Classroom effect · Self-concept of ability · Academic skills · Primary school

Students' achievement-related beliefs play an important role in school by directing behavior and effort in learning situations (e.g., Atkinson 1964; Bandura 1997; Eccles et al. 1983; Wigfield et al. 2006). Students who believe that they can and will do well at school are more likely to perform better and to engage in an adaptive manner in academic tasks than students who have a negative self-perception and expect to fail (Pintrich and Schunk 2002). Several

✉ Kati Vasalampi
kati.vasalampi@jyu.fi

¹ Department of Psychology, University of Jyväskylä, P. O. Box 35, 40014 Jyväskylä, Finland

² Department of Teacher Education, University of Jyväskylä, P. O. Box 35, 40014 Jyväskylä, Finland

³ School of Educational Sciences and Psychology, Philosophical Faculty, University of Eastern Finland, P. O. Box 111, 80101 Joensuu, Finland

concepts pertaining to achievement-related beliefs have been introduced, such as perceived competence (Harter 1982), academic self-concept (Shavelson et al. 1976), and self-concept of ability (e.g., Nurmi and Aunola 2005), the last of which will be used in the present study.

Despite certain differences in emphasis, these concepts all refer to students' own understanding of their abilities in academic situations (for a review, see Bong and Skaalvik 2003). Furthermore, they all emphasize that the understanding of the ability is based not only on an individual's own performance but also on social comparison with the reference group (for empirical evidence, see, e.g., Chiu et al. 2017). The big-fish-little-pond effect (BFLP effect) model (Marsh et al. 2007) argues that students compare their own academic ability with the academic abilities of their classmates and use this social comparison impression as a basis for forming their self-concept of ability. Previously, this model has primarily been tested among adolescents and in selective school systems (Marsh et al. 2007, Marsh et al. 2018); however, the model has rarely been tested among younger students (see Guo et al. 2018), although primary school age might be the most critical period for forming of self-concept of ability. In the present study, we examine whether the effects indicated by the BFLP effect model can already be found in the primary grades and in a school system that does not select and track students according to different abilities, that is, in the Finnish educational context.

The big-fish-little-pond effect model

Individual academic performance has been shown to provide the basis for the development of students' self-concept of abilities in primary school (Aunola et al. 2002; Chapman and Tunmer 1997). During the first school years, students acquire experiences of diverse academic tasks and related feedback on their success. Increasing cognitive maturation allows the students to make more accurate evaluations of this feedback and to compare their abilities and success to those of their classmates. During this time, the associations between students' academic performance and self-concept of ability increase in intensity over their previous state (Denissen et al. 2007; Eccles et al. 1998).

The skill-development model argues that academic performance is important in the formation of self-concept among younger students (Calsyn and Kenny 1997). In support of this developmental view, Aunola et al. (2002) showed that good reading skills in grade 1 (7-year-olds) make a positive contribution to later self-concept of ability, while poor reading skills constitute a risk factor for the development of negative self-concept of ability. Chapman and Tunmer (1997) demonstrated that children's pre-reading performance at age 5 was an important predictor of their reading-related self-concept at age 7. Similarly, in a longitudinal study assessing both students' self-concept of ability and performance in reading and mathematics during grades 1, 2, 4, and 7, Viljaranta et al. (2014) showed that performance during previous grades predicted students' subsequent self-concept of ability. However, gender may moderate this relation between performance and self-concept. It has been shown, for instance, that, while boys and girls perform equally well in mathematics and in literacy, gender differences in self-concept exist from quite a young age, with boys holding more positive self-concept for mathematics and girls holding more positive self-concept for literacy (Denissen et al. 2007; Eccles et al. 1993; Wouters et al. 2013).

In addition to children's own academic performance, the performance level of their classmates may also have an effect on students' self-concept of ability. It has been suggested that equally skilled students have lower self-concept of ability when they are placed in a

classroom where the average performance level is high than when they attend a low-performing classroom (Marsh 1987; Marsh et al. 2018). Thus, the same absolute level of performance can lead to different self-concept depending on the contexts in which students evaluate themselves. This phenomenon—known as *the big-fish-little-pond effect* (Marsh et al. 2007, 2008)—has been shown to be highly specific to the academic component of self-concept. For example, Marsh (1987) demonstrated that there was a large negative classroom effect on academic self-concept but little or no negative classroom effect on general self-concept or self-esteem.

In statistical terms, BFLP effect has been operationalized in terms of a concurrent path model where both students' own performance and the performance level of the classroom predict individual students' self-concept of ability (see Fig. 1). The effect of students' own performance on self-concept is assumed to be positive (e.g., the better the skills of a student in mathematics, the higher the student's self-concept in mathematics), whereas the effect of the performance level of the class is assumed to be negative (the better the skills of the class in mathematics, the lower the individual student's self-concept in mathematics—after accounting for the student's own skills in mathematics).

The mechanism underlying the BFLP effect has been explained from a number of different perspectives (Marsh and O'Mara 2010), which all emphasize that comparison processes are inevitable in a classroom (Marsh et al. 2007, 2018). The negative effect has been shown to be similar regardless of students' own performance level (Marsh and Hau 2003). In samples involving adolescents and young adults in selective school systems, evidence has indicated cross-cultural generalizability of the BFLP effect model (Marsh and Hau 2003) and also support for the model in a variety of academic domains (Parker et al. 2013). Altogether, these studies attest that evidence for the BFLP effect is highly robust among adolescents (Marsh et al. 2018).

Nonetheless, the previous research on BFLP effect has at least three limitations. First, the BFLP effect model has mostly been applied in the analysis of the effects of classes or schools on adolescents' or young adults' academic self-concept (e.g., Jonkmann et al. 2012; Marsh and O'Mara 2010; Parker et al. 2013; Trautwein et al. 2006). There is surprisingly scant research testing the BFLP effect among younger primary school students, even though examination of the BFLP effect is particularly interesting with respect to these years when students' self-perception of their academic performance is being formed (see Guo et al. 2018; Marsh et al. 2015; Pinxten et al. 2015; Thijs et al. 2010; Trautwein et al. 2008; Wouters et al. 2013).

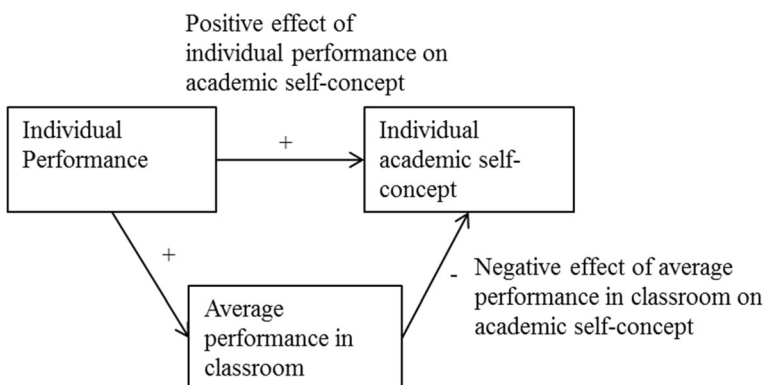


Fig. 1 The big-fish-little-pond effect (BFLP effect) model (adapted from Marsh et al. 2007)

Trautwein et al. (2008) examined the BFLP effect among 10-year-old children, but they focused on physical activity rather than academic skills. Thijs et al. (2010) and Wouters et al. (2013) studied the BFLP effect among Dutch and Belgian primary school students, finding that students in the last grade of primary schools ($M_{\text{age}} = 11\text{--}12$ years) used their classmates as referents in assessing their own academic performance. In a Belgian study by Pinxten et al. (2015), strong negative correlations were documented between class-average performance in math and literacy and academic self-concept in these subjects already in grade 4. Marsh et al. (2015) investigated evidence for negative classroom effect on students' self-concept in mathematics by using data from fourth and eighth grade students in three Middle Eastern Islamic countries (Iran, Kuwait, and Tunisia), four Asian countries (Hong Kong, Japan, Taiwan, and Singapore), and six Western countries (Australia, UK, Italy, Norway, Scotland, and the USA). They documented a positive effect of students' performance and a negative effect of class average performance on self-concept in mathematics in all countries in both age cohorts. When focusing on the results from Western countries, the variation in the negative BFLP effect was notably wide among fourth graders: it ranged from -0.134 in Norway to -0.482 in Italy. A study by Guo et al. (2018), which also focused on the BFLP effect among fourth graders, indicated that the BFLP effect might be weaker in reading than in math.

While there is currently some evidence that supports the assumption of negative BFLP effect already in the primary school context, more research is needed to establish whether these findings could be replicated in other educational systems. Furthermore, previous findings by Marsh et al. (2015) suggest replications of the BFLP effect would be especially relevant in Nordic countries where the effect might be different if compared to other Western countries. Thus, in the present study, data from students in grade 3 (age 9), grade 4 (age 10), and grade 6 (age 12) were used to examine whether the assumptions of the BFLP effect model already hold during the primary school years in Finland. Because the literature indicates slightly different effects for literacy and math (Guo et al. 2018), both subject domains were included in the study.

Second, prior studies examining BFLP effects on students' self-concept have typically been conducted in academically selective school systems (e.g., Jonkmann et al. 2012; Marsh et al. 2007; Marsh and O'Mara 2010; Parker et al. 2013; Schurtz et al. Shurtz et al. 2014; Trautwein et al. 2006), where students are tracked according to their performance level. These studies have shown a particularly strong negative effect of classroom performance among all students attending academically selective schools or classes (Marsh and Hau 2003). The BFLP effect model has seldom been studied in nonselective educational settings, and more specifically, only a few researchers (Guo et al. 2018; Marsh et al. 2015; Pinxten et al. 2015) have studied it in nonselective primary school systems, in which the present study was conducted.

Finally, although it has been claimed that the BFLP effect is robust, universal, and scarcely affected by any moderators (cf. Marsh et al. 2008), gender differences in the BFLP effect have only seldom been systematically examined (Marsh et al. 2007; Plieninger and Dickhäuser 2015; Preckel et al. 2008; Thijs et al. Thijs et al. 2010), and the findings for gender have been contradictory. For example, Marsh et al. (2007) concluded that difference between girls and boys in BFLP effect for math self-concept is marginal. In contrast, results by Plieninger and Dickhäuser (2015) indicated that gender moderates the BFLP effect, as the negative classroom effect in science was substantially larger for girls than for boys. Further, Preckel et al. (2008) found that the negative classroom effect was particularly large among girls attending a special class for gifted. However, given that only a few researchers have empirically investigated the

generalizability of the BFLP effect across gender, and of those only Thijs et al. (Thijs et al. 2010) have focused on primary school, further research is needed on the topic.

Finnish school system

In Finland, compulsory education beginning at age 7 is an integrated 9-year structure intended for the entire age group. Finnish society strongly emphasizes the importance of equal educational opportunities for everyone, and, as a result, most Finnish children attend the public school system (only 1.4% attend private schools), and compulsory education is provided for all students completely free of charge (including free learning materials, books, and lunch). Furthermore, primary schools do not select or test students as part of admission (except for some schools with special emphasis), but instead students are from the catchment area (i.e., from neighborhood) and they are designated a place by the municipality educational authorities. An overwhelming majority of students are enrolled in their nearest school. The classes are formed by the principals more or less randomly instead of assigning students to classes based on ability for instance. Parents can, however, make a wish, that their child could be enrolled in the same class with his or her best friend, but this is not guaranteed. Students often continue with same classmates for 6 years until the end of the primary grades.

During the primary school phase, grades 1–6, students typically are taught all or most of the subjects by their class teacher. Class teachers in Finland are required to have a university Master of Education degree. Teachers are responsible for autonomously applying the curriculum guidelines and assessing student progress on the basis of the objectives of the national core curriculum, and no national tests or evaluations are administered in primary school. All students have the right to general and special support, individual guidance, and tailored instruction at school if needed (National Core Curriculum 2014).

Present study

The aim of this study is to test whether the assumptions of the BFLP effect model hold for primary school students in a nonselective educational system. First, we set out to examine to what extent students' performance in literacy and in mathematics, assessed at grades 3, 4, and 6, is related to their self-concept in these subjects. In line with previous studies, we expected (Hypothesis 1) that academic performance in a particular subject would predict the level of self-concept in that subject (Aunola et al. 2002; Chapman and Tunmer 1997; Viljaranta et al. 2014). Second, we examine to what extent the average performance level in the classroom is related to students' self-concept in literacy and in mathematics at grades 3, 4, and 6. The findings of the previous studies have demonstrated a classroom effect during primary grades in certain subjects (Guo et al. 2018; Marsh et al. 2015; Pinxten et al. 2015; Thijs et al. 2010; Trautwein et al. 2008). Thus, we hypothesized (Hypothesis 2) that classroom average performance level would be negatively related to students' self-concept in literacy and in mathematics after controlling for individual student performance (Marsh et al. 2008), but we hypothesized that the effect might be stronger for math than for literacy (see Guo et al. 2018). Third, because previous research has found contradictory findings concerning the differences between girls and boys with respect to negative classroom effects (e.g., Marsh et al. 2008; Plieninger and Dickhäuser 2015; Preckel et al. 2008), we examined gender differences in these associations. However, no exact hypotheses for gender differences were set.

Method

Participants and procedure

This study is part of an extensive longitudinal First Steps study (Lerkkanen et al. 2006-2016) comprising approximately 2000 students who in the first phase were followed up from the beginning of their kindergarten year (year 2006) to the end of the sixth grade (year 2012), with simultaneous data gathering from their parents and teachers. The sample was drawn from four municipalities in different parts of Finland. In three of these municipalities, the whole age cohort participated, while in the fourth municipality, the participating students comprised approximately half of the age cohort. Parents gave their written consent for their children's participation in the study.

The participants whose data were used in the present analyses consisted of an intensively followed subsample of 504 students (237 girls and 267 boys) from 133 grade 3 classrooms (the average age was 8 years and 3 months) attending general education (those attending special education classes were excluded from the present data). The selection of the subsample was performed in a stratified fashion so that there was at least one child from each classroom. This subsample consisted of children identified at the end of kindergarten as at risk for reading difficulties because of lower level of pre-reading skills ($n = 214$), as well as randomly selected control children from the same classrooms ($n = 290$). Because of the high number of at-risk children, we compared the subsample ($n = 504$) to the total sample ($N = 2000$) of the follow-up in terms of children's math and reading skills. The results of the independent samples *t*-test indicated that the present subsample did not differ from the total sample in terms of academic skills. Of these 504 grade 3 students, data was available for 482 students (225 girls and 257 boys from 130 classrooms) in grade 4 and for 365 students (162 girls and 203 boys from 119 classrooms) in grade 6.

Students were asked to provide ratings in individual test situations regarding their self-concept of ability in literacy and mathematics in the spring of grades 3, 4, and 6. In addition, they were tested at the same time points on their literacy and mathematics skills in group test situations with their classmates. Calculation of the classroom average performance scores for the BFLP effect models (i.e., aggregated performance) was based on information regarding all students from each classroom who participated in the group tests ($N = 1889$ at grade 3; $N = 1852$ at grade 4; and $N = 1769$ at grade 6).

In terms of family background, 7% of the children's parents did not have education beyond comprehensive school (compulsory education up to the completion of grade 9), 31% had completed upper-secondary education (senior high school or vocational school, grades 10–12), 36% had a bachelor's degree or a vocational college degree (3-year education at a college or university), and 26% had a master's degree or higher (i.e., licentiate or doctoral degree). The sample was fairly representative of the Finnish population (Statistics Finland 2007).

Measures

Self-concept in literacy and mathematics Students' self-concept in literacy (reading and spelling) and in mathematics was measured using the Self-Concept of Ability Scale (Nicholls 1978; for validity, see Aunola et al. 2002). In the test, students were presented with 10 circles arranged in a column from the top to the bottom of the page. The students were told that the chain of circles represents a group of students, with the topmost circle (value 1) representing

the student who is best at a certain subject, the circle at the bottom (value 10) representing the student with the poorest skills, and the others representing students with skills in between. The participants responded by pointing to one of the circles. For the analyses, the scale was reversed (i.e., value 10 standing for high self-concept of ability and value 1 standing for low self-concept of ability). In grade 3, the question was similar for literacy and for mathematics: "Can you show me how good you are at reading and spelling/at mathematics? Which one is you?". In grades 4 and 6, the variable measuring students' self-concept in literacy was constructed based on six questions. Three of the questions measured self-concept of ability in reading (general reading-related self-concept, reading fluency, and reading comprehension): "Can you show me how good you are at reading/reading fast/understanding what you read? Which one is you?". Three of the questions measured self-concept of ability in spelling (general writing-related self-concept, spelling, and productive writing): "Can you show me how good you are at writing/spelling/productive writing? Which one is you?". Cronbach alpha reliability was 0.81 at grade 4 and 0.89 at grade 6. The variable measuring grade 4 and grade 6 students' self-concept in mathematics was constructed using three questions measuring self-concept of ability in mathematics (general math skills, automatization of adding and subtraction skills, and multiplication): "Can you show me how good you are at math/multiplication/mental arithmetic? Which one is you?". Cronbach alpha reliability was 0.84 in grade 4 and 0.88 in grade 6.

Literacy performance A variable measuring students' literacy performance was constructed based on students' scores in group tests of reading fluency, reading comprehension, and pseudoword spelling.

Reading fluency was assessed using a group-administered subtest of the nationally standardized reading test battery (Lindeman 1998). In this speed test, a maximum of 80 items could be attempted in a 2-min time limit. For each item, there was a picture with four words next to it, and students were asked to read silently the four (phonologically similar) words and to then draw a line connecting the picture to the word semantically matching it. The measure was scored by calculating the number of correct answers (maximum score possible = 80). Because of the nature of this speed test, the score reflects both the student's fluency in reading the stimulus words and his or her accuracy in making the correct choice from among the alternatives. In the highly transparent orthography of the Finnish language, differences between students have typically been identified using this type of reading test (Holopainen et al. 2001).

Reading comprehension was assessed using a group-administered subtest of the nationally normed reading test battery (Lindeman 1998). The students were asked to silently read a fictional story and then answer 12 multiple-choice questions. The students received one point for each correct answer (maximum score possible = 12). The students completed the task at their own pace, but the maximum time allotted was 45 min.

Pseudoword spelling was assessed using a group-administered task with 8 items consisting of a one-syllable pseudoword (1 item: vuil), two-syllable pseudowords (2 items: saihdi; raalsku), and three- and four-syllable pseudowords (5 items: hiuruutti; seivolssi; paunitteri; ruustivaimu; nuppasengit) (Torppa et al. 2010). Each item was presented orally twice. The score was the number of correctly spelled items (maximum score = 8). The composite score for literacy performance in grades 3, 4, and 6 was formed by calculating a mean based on the standardized test scores of reading fluency, reading comprehension, and pseudoword spelling.

Mathematics performance A variable measuring students' mathematics performance was constructed in grades 3 and 4 based on students' scores in fluency in arithmetic skills and arithmetic reasoning. Each assessment was weighed as the same in computing the mean of mathematics performance. In grade 6, students' performance in mathematics was based on their scores in fluency in arithmetic skills. Thus, the focus on the assessments employed was on mastery of basic arithmetic skills.

Fluency in arithmetic skills was assessed using the group-administered Basic Arithmetic Test (Aunola and Räsänen 2007; see also Räsänen et al. 2009). In this speed test, a maximum of 28 items, consisting of 14 items for addition (e.g., $2 + 1 = ?$ and $3 + 4 + 6 = ?$) and 14 items for subtraction (e.g., $4 - 1 = ?$ and $20 - 2 - 4 = ?$), is attempted in a 3-min time limit (in grade 6, four of these items included also division or multiplication). Task difficulty increases gradually across the test. The test indexes a combination of speed and accuracy (see Zhang et al. 2014). The final score is the total number of correct answers (maximum score possible = 28).

Arithmetic reasoning was assessed using a NMART test (Koponen and Räsänen 2003; see also Langdon and Warrington 1997) with 30 items in which students are asked to continue a series of three numbers by adding a fourth number that would complete the series. For example, an item would involve a series of three numbers (e.g., 3, 5, 7), and the student needs to circle the number out of the four additional numbers given as alternatives that best fits as the fourth number in the series. After four practice items, the students completed the test at their own pace, but the maximum time allotted was 10 min for 30 tasks. It was assumed that on average students would complete the tasks approximately in 5 min, and therefore, the test was not a speed test. One point was given for each correct answer. The final score was the total number of correct answers (maximum score possible = 30).

The composite score for mathematics performance in grades 3 and 4 was formed by calculating a mean based on the standardized test scores of arithmetic fluency and arithmetic reasoning. In grade 6, fluency in arithmetic skills was used as the measure of mathematics performance.

Classroom average performance Students' average performance in literacy (reading fluency, reading comprehension, and pseudoword spelling) was aggregated across the students in a classroom, i.e., students' mean score of performance level was calculated for each classroom. Students' performance in mathematics (arithmetic and arithmetic reasoning) was also aggregated across students in a classroom by calculating mean score of the performance level in mathematics in each classroom.

Analytical strategy

The analyses were performed using the Mplus statistical package (version 7.01; Muthén and Muthén 1998–2012). The standard missing at random (MAR) approach was applied (Muthén and Muthén 1998–2012). This missing-data method uses all the data that is available in order to estimate the model without imputing data. The parameters of the models were estimated using the full information maximum likelihood (FIML) estimation with non-normality robust standard errors (MLR estimator). As the data were hierarchical in nature (i.e., students were nested in classrooms), we calculated intraclass correlations (ICCs) and design effects. The ICCs for self-concept of ability in literacy were 0.02, 0.01, and 0.01, and p values = 0.78, 0.81,

and 0.81, in grades 3, 4, and 6, respectively. The ICCs for self-concept of ability in math were 0.03, 0.01, and 0.01, and p values = 0.87, 0.89, and 0.88, in grades 3, 4, and 6, respectively. The design effects were for self-concept of ability in literacy, 1.06, 1.03, and 1.02, and for self-concept of ability in math, 1.08, 1.03, and 1.02, respectively. Hox and Maas (2002) suggested that analyzing multilevel data as single-level data can yield acceptable (not overly biased) parameter estimates and inferential tests, if the design effects are smaller than 2.0. Therefore, we used the Type = COMPLEX approach (Muthén and Muthén 1998–2012). This COMPLEX approach estimates the model at the level of the whole sample, but corrects for distortions in standard errors in estimation caused by the clustering of observations (i.e., clustering effect of students nested in classrooms). Such methodology has been used previously by Marsh and O'Mara (2010) when investigating the BFLP effect. Differences between girls and boys in the BFLP effect model were tested using a multigroup approach, and differences between models were tested with the Satorra–Bentler scaled χ^2 test for difference. The goodness-of-fit of the estimated models was evaluated using the following four indicators: χ^2 test, comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR).

Results

Descriptive analyses

Descriptive statistics, correlations between the observed variables, and mean differences between girls and boys in performance and self-concept in literacy and mathematics are presented in Tables 1, 2, and 3. Students' performance in literacy and their self-concept in literacy, as well as performance in math and self-concept in math, had a statistically significant positive correlation in grades 3, 4 and 6. Independent samples t test analysis showed that literacy performance and self-concept in literacy were both significantly higher for girls than for boys across all grades. In mathematics, performance of boys was significantly better than that of girls' in grades 3 and 4, but in grade 6, girls and boys did not differ in mathematics performance. However, self-concept in mathematics was significantly higher for boys than for girls across all grades.

Models for the big-fish-little-pond effect

The path models testing the assumptions of the BFLP effect model examined the associations between students' individual academic performance and self-concept of ability, as well as the associations between classroom average performance and self-concept of ability. In this model, evidence of a negative classroom effect would be indicative of the BFLP effect. Models were calculated separately for literacy and mathematics within each grade level. Differences between girls and boys in the BFLP effect models were tested using a multigroup approach.

Path models in grade 3 The results of grade 3 models are presented in Fig. 2. For literacy, the gender-constrained model (i.e., a similar model for girls and boys) suggested that there is a difference between girls and boys [$\chi^2(3) = 8.10, p = 0.04$; CFI = 0.97, RMSEA = 0.08, SRMR = 0.08]. The Satorra–Bentler scaled χ^2 test for difference [$\Delta\chi^2(\Delta 1) = 9.89, p = 0.002$] indicated that individual performance was more strongly related to self-concept in literacy among girls

Table 1 Correlations, means (*M*), standard deviations (SD), and mean differences between girls and boys (*t*) at grade 3

	1.	2.	3.	4.	5.	6.	<i>M</i> ^{girls}	<i>SD</i> ^{girls}
1. Individual literacy performance ^a	1.00	0.35***	0.37***	0.17*	0.50***	0.19**	0.11	0.55
2. Individual math performance ^b	0.47***	1.00	0.23***	0.43***	0.22***	0.37***	-0.14	0.76
3. Classroom literacy performance	0.35***	0.16*	1.00	0.43***	0.11	0.08	0.16	0.45
4. Classroom math performance	0.22***	0.40***	0.53***	1.00	-0.02	0.10	0.12	0.64
5. Self-concept in literacy	0.30***	0.17**	0.02	-0.02	1.00	0.39***	6.94	1.50
6. Self-concept in math	0.14*	0.37***	-0.08	-0.02	0.24**	1.00	6.58	1.92
<i>M</i> ^{boys}	-0.17	0.04	0.06	0.05	6.35	7.35		
<i>SD</i> ^{boys}	0.68	0.94	0.53	0.77	1.72	1.83		
Mean differences between genders (<i>t</i>)	5.08***	-2.38*			3.92***	-4.60***		

Note. *n* = 504 (237 girls and 267 boys), but due to missing data, the values in the tables were not always computed on the whole groups of girls and boys. Values for girls are above the diagonal and for boys are below the diagonal

p* < 0.05; *p* < 0.01; ****p* < 0.001

^a The composite of the students' scores on the reading fluency, reading comprehension, and pseudoword spelling tests

^b The composite of the students' scores on the fluency in arithmetic skills and arithmetic reasoning tests

than boys. The analyses did not show differences between genders concerning the negative effect of classroom performance in literacy, and the effect of average classroom performance was nonsignificant for both girls and boys.

For mathematics, the gender-constrained model (i.e., similar model for girls and boys) provided an excellent fit [$\chi^2(3) = 2.22, p > 0.05; CFI = 1.00, RMSEA = 0.00, SRMR = 0.03$]

Table 2 Correlations, means (*M*), standard deviations (SD), and mean differences between girls and boys (*t*) at grade 4

	1.	2.	3.	4.	5.	6.	<i>M</i> ^{girls}	<i>SD</i> ^{girls}
1. Individual literacy performance ^a	1.00	0.52***	0.40***	0.19*	0.25**	0.31***	0.13	0.60
2. Individual math performance ^b	0.55***	1.00	0.22**	0.35***	0.16*	0.58***	-0.13	0.73
3. Classroom literacy performance	0.46***	0.29***	1.00	0.53***	0.00	0.08	0.13	0.54
4. Classroom math performance	0.32**	0.43***	0.66***	1.00	-0.01	0.16*	0.08	0.66
5. Self-concept in literacy	0.42***	0.18*	0.16	0.06	1.00	0.49***	7.20	1.21
6. Self-concept in math	0.31***	0.48***	0.11	0.12*	0.46***	1.00	6.19	1.80
<i>M</i> ^{boys}	-0.15	0.06	0.06	0.02	6.82	7.12		
<i>SD</i> ^{boys}	0.68	0.92	0.72	0.85	1.20	1.58		
Mean differences between genders (<i>t</i>)	4.75***	-2.57*			3.48**	-5.88***		

Note. *n* = 487 (231 girls and 256 boys), but due to missing data, the values in the tables were not always computed on the whole groups of girls and boys. Values for girls are above the diagonal and for boys are below the diagonal

p* < 0.05; *p* < 0.01; ****p* < 0.001

^a The composite of the students' scores on the reading fluency, reading comprehension, and pseudoword spelling tests

^b The composite of the students' scores on the fluency in arithmetic skills and arithmetic reasoning tests

Table 3 Correlations, means (*M*), standard deviations (SD), and mean differences between girls and boys (*t*) at grade 6

	1.	2.	3.	4.	5.	6.	<i>M</i> ^{girls}	<i>SD</i> ^{girls}
1. Individual literacy performance ^a	1.00	0.41***	0.36***	0.11	0.32***	0.18*	0.16	0.59
2. Individual math performance	0.47***	1.00	0.09	0.37***	0.26**	0.56***	15.72	3.59
3. Classroom literacy performance	0.41***	0.17*	1.00	0.36***	0.10	0.03	0.16	0.58
4. Classroom math performance	0.14*	0.34***	0.56***	1.00	0.01	0.08	16.42	1.39
5. Self-concept in literacy	0.26**	0.17*	-0.08	-0.12	1.00	0.38***	7.01	1.23
6. Self-concept in math	0.27***	0.49***	-0.01	0.05	0.50***	1.00	6.01	1.81
<i>M</i> ^{boys}	-0.17	16.10	-0.02	16.22	6.43	6.91		
<i>SD</i> ^{boys}	0.67	3.62	0.71	1.53	1.56	1.85		
Mean differences between genders (<i>t</i>)	4.95***	-0.97			3.83***	-4.65***		
		ns						

Note. *n* = 365 (162 girls and 203 boys), but due to missing data, the values in the tables were not always computed on the whole groups of girls and boys. Values for girls are above the diagonal and for boys are below the diagonal

p* < 0.05; *p* < 0.01; ****p* < 0.001

^a The composite of the students' scores on the reading fluency, reading comprehension, and pseudoword spelling tests

(Fig. 2), and further analyses showed that girls and boys did not differ statistically in any path. Individual performance in mathematics was significantly and positively related to both girls' and boys' self-concept in mathematics. Although the χ^2 test for difference [$\Delta\chi^2(\Delta 1) = 1.69, p > 0.05$] indicated that the negative classroom effect in mathematics was similar among boys and girls, the path coefficients indicated a significant negative classroom effect for boys but a nonsignificant effect for girls.

Taken together, the grade 3 models indicated that high-level individual performance was positively related to students' self-concept in literacy and in mathematics, but in literacy, the relation was stronger among girls than boys. Average classroom performance level was negatively related only to boys' self-concept in mathematics.

Path models in grade 4 Next, we tested the effects of individual performance and classroom average performance for literacy and mathematics in grade 4 (Fig. 3). The gender-constrained model provided an excellent fit for the literacy model: [$\chi^2(3) = 3.70, p > 0.05; CFI = 1.00,$

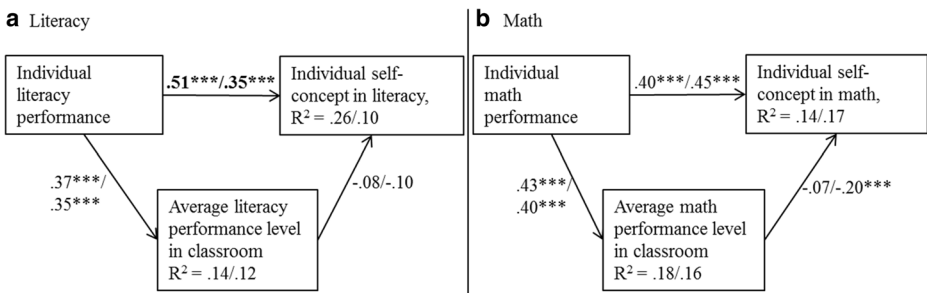


Fig. 2 Effect of individual performance and average classroom performance on students' self-concept in literacy (a) and in mathematics (b) in grade 3: standardized solution. Note. ****p* < 0.001. The first value is for girls and the second is for boys (significantly different values are bolded)

RMSEA = 0.03, SRMR = 0.06], and further analyses indicated that girls and boys did not differ statistically in any path. A significant positive association was found between students' individual performance in literacy and their self-concept in literacy for both girls and boys. Classroom average performance was not significantly associated with either girls' or boys' self-concept in literacy.

For mathematics, the gender-constrained model provided a poor fit [$\chi^2(3) = 10.31, p = 0.02; CFI = 0.97, RMSEA = 0.10, SRMR = 0.11$], indicating that certain differences between girls and boys were present. The Satorra–Bentler scaled χ^2 test for difference [$\Delta\chi^2(\Delta 1) = 9.98, p = 0.002$] showed that a path from individual performance to self-concept in mathematics was stronger among girls than boys. Although the chi-squared difference test for difference [$\Delta\chi^2(\Delta 1) = 1.96, p > 0.05$] failed to show a difference in the classroom effect between girls and boys, the path coefficients indicated a significant negative classroom effect for boys but not for girls.

Taken together, the grade 4 models indicated that individual performance was positively related to students' self-concept in literacy and in mathematics. Furthermore, in mathematics, this association was stronger for girls than for boys. However, as in grade 3, the assumed negative BFLP effect was found only for boys and only in mathematics.

Path models in grade 6 Finally, we tested the effects of individual performance and classroom average performance in grade 6 (Fig. 4). The gender-constrained model provided an excellent fit for the literacy model: [$\chi^2(3) = 4.80, p > .05; CFI = 0.98, RMSEA = 0.06, SRMR = 0.04$]. The results did not indicate differences between girls and boys in the relation between individual performance in literacy and self-concept in literacy: Individual performance in literacy was significantly and positively related to both girls' and boys' self-concept in literacy. The findings concerning average classroom performance indicated a negative classroom effect [$\Delta\chi^2(\Delta 1) = 3.98, p = .045$] for boys but not for girls in literacy.

For mathematics, the gender-constrained model provided an excellent fit [$\chi^2(3) = 0.48, p > .05; CFI = 1.00, RMSEA = 0.00, SRMR = 0.03$]. Furthermore, the χ^2 -test for difference indicated no differences between genders. The results showed a strong positive relation between individual performance and self-concept in mathematics for girls as well as for boys. The negative classroom effect was significant for both girls and boys.

In sum, the findings for Grade 6 indicated that the positive relation between individual performance and students' self-concept in literacy and mathematics was strong and similar

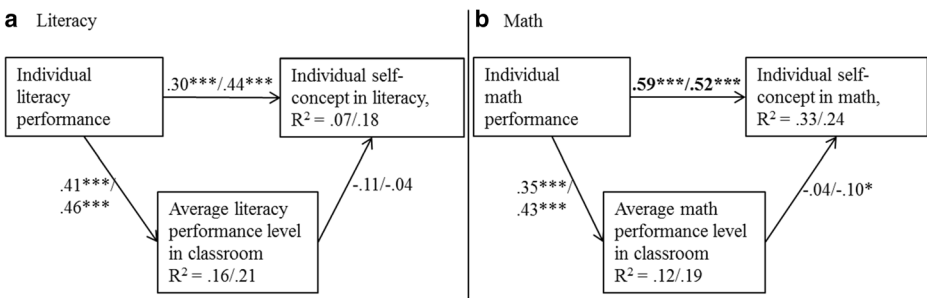


Fig. 3 Effect of individual performance and average classroom performance on students' self-concept in literacy (a) and in mathematics (b) in grade 4: standardized solution. Note. * $p < 0.05$; *** $p < 0.001$. The first value is for girls and the second is for boys (significantly different values are bolded)

among girls and boys. Negative classroom effect was present in literacy only for boys, but in mathematics this effect was present regardless of gender.

Discussion

According to the BFLP effect model, a student's self-concept of ability is dependent on the performance level of the group, as well as the student's own individual performance level (Marsh et al. 2007). In the present study, using a Finnish sample, we examined whether the assumptions of the BFLP effect model would already be visible among primary school students (Grades 3, 4, and 6) in a nonselective school system. Furthermore, we examined whether there are gender differences in the BFLP effect. The results showed that students' individual performance in both literacy and in mathematics was positively associated with their self-concepts in literacy and mathematics across all grades. A small negative classroom effect proposed by the BFLP effect model emerged in mathematics. This was indicated by lower self-concept of ability in classrooms with high average performance level among boys in all grades, whereas among girls this was found only in Grade 6. In literacy, the data did not show evidence for the negative classroom effect proposed by the BFLP effect model in Grades 3 or 4, and in Grade 6 it was found only among boys. The finding of relatively small negative classroom effects in comparison to some prior studies may suggest that classrooms void of selection and tracking may protect against comparisons in classrooms which typically have a decreasing impact on students' self-concept of ability.

The results of the present study showed in line with previous literature that students' own performance level was positively associated with their self-concept of ability at each grade level. This was true for both literacy and mathematics and for both boys and girls: the better the individuals' own performance level, the higher the level of their self-concept of ability in that particular subject. The third graders' performance in literacy and mathematics was already highly related to their self-ratings of their standing in comparison to their peers in the respective subjects. This result is in line with previous studies showing a positive association between performance and academic self-concept (Aunola et al. 2002; Chapman and Tunmer 1997; Viljaranta et al. 2014). As suggested by motivational models, feedback from teacher and peers, along with skill level and accomplishments, is likely to affect students' self-concept of ability (for a review see Muenks et al. 2018). Teacher's feedback in the classroom, in

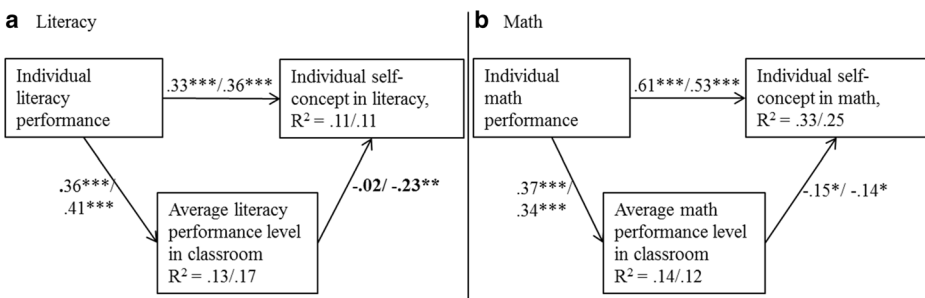


Fig. 4 Effect of individual performance and average classroom performance on students' self-concept in literacy (a) and in mathematics (b) in grade 6: standardized solution. Note. $*p < 0.05$; $**p < 0.01$; $***p < 0.001$. The first value is for girls and the second is for boys (significantly different values are bolded)

particular, is seen to contribute to students' evolving understanding of their abilities in academic situations (Burnett 2003; Furtak et al. 2016; Pesu et al. 2016).

The results provided some evidence for the assumed negative classroom effect of self-concept of ability in mathematics, especially for boys after controlling for individual student performance, whereas evidence for a negative classroom effect in literacy was very scant and emerged only in Grade 6 and only for boys. In mathematics, the average performance level in the classroom was negatively associated with boys' self-concept of ability in mathematics at each grade level, although the effect was small in comparison to recent result from Belgian primary school system (see Pinxten et al. 2015). For girls, the evidence failed to show a negative classroom effect in mathematics in Grades 3 and 4, but a small effect was found in Grade 6. In literacy, the analyses failed to show a negative classroom effect for self-concept in Grades 3 or 4, and in Grade 6 the effect was found only for boys. The findings thus provide some support for the BFLP effect model which proposes that social comparison processes affect students' academic self-concept already in primary school classrooms (see also Guo et al. 2018; Marsh et al. 2015; Pinxten et al. 2015; Thijs et al. 2010). However, given the small effect size, the support was weak.

It seems that in the Finnish nonselective school system, where children's performance level in classroom can be quite heterogeneous due to lack of selection or tracking, a negative classroom effect on student's self-concept of ability is small. This suggestion is in line with results by Chiu et al. (2017) showing that high heterogeneity in primary school classrooms (in terms of family background and past achievement) benefit students reading achievement. The findings also corroborate the results by (Marsh et al. 2007, 2018) showing that in Norway (where the educational system is very similar to Finland), the effect of negative BFLP effect among fourth graders was also small in comparison to several other Western countries. However, it is important to note, that BFLP effect has rarely been examined among young primary school students. It is possible that the scant evidence for BFLP effect in the present study is also partially due to participants' age. Nicholls (1978) argued already in late seventies that children in between 9 to 12 years of age can differentiate ability and effort as causes of outcomes, but do not always apply this understanding. This statement has been criticized recently, for example by Cimpian (2017; for a review see Moenks et al. 2018) who showed that children are able to understand abilities at a much earlier age. Thus, future research should continue to examine the developmental differentiation of self-concepts of ability and the role of the context on BFLP effect in primary school.

The results in mathematics differed from those in literacy, which is in accordance with recent findings by Guo et al. (2018). The difference between the subjects may be associated with the different nature of these subjects, the beliefs of the capabilities required to succeed in them, and the visibility of differences in skill levels. For negative BFLP effect to occur, students need to have accurate information about the accomplishments of their classmates. Students' exposure to comparable information may be more likely in mathematics than in literacy. Mathematics is a hierarchically structured content and skill area, where new knowledge is cumulatively based on previously established routines and understanding. In mathematics, classmates may more easily notice when a student falls behind the rest of the class. In literacy, by contrast, comparing literacy skills after the basic reading acquisition stage is not as clear as in mathematics, especially in the Finnish context where children typically reach fluent and accurate reading skills by the end of second grade (Lyytinen et al. 2006). Another possibility is, as Guo et al. (2018) suggest, that the BFLP effect is stronger when students' opportunities to evaluate their competence is restricted to school environment. One's level of literacy skills may be more often

than math skills be used as a resource and manifested in the daily life outside of school, and, thus, students may construct their self-concept in literacy using a broader context and experiences gathered from formal and informal settings than that for math.

The findings revealed interesting differences between girls and boys in self-concept of ability, and, particularly, in relation to classroom effects. The analyses showed that girls performed better in literacy at all grade levels and that their literacy self-concept was higher than that of boys. Moreover, negative BFLP effect did not emerge in literacy among girls. By contrast, boys performed better than girls in mathematics in Grades 3 and 4, and their mathematics self-concept was higher than that of girls in all grades. However, among boys, small negative classroom effect was present in mathematics already in Grade 3 and continued to be evident in Grades 4 and 6, whereas among girls it emerged only in Grade 6. These findings extend our understanding of BFLP effect, first, by showing that average classroom performance has different effect on girls' and boys' self-concept of ability (see also Plieninger and Dickhäuser 2015; Preckel et al. 2008), and that the effect is dependent on the school subject. The results imply that boys start to compare their performance with their classmates earlier than girls and that the classroom context affects boys' mathematics self-concept negatively already at an early phase in their school career.

The results highlight the relevant role of the classroom environment and the potential influences of teacher-student and peer relations and of practices of evaluation and feedback on students' self-concept of ability. Comparison with peers and apprehension about one's own standing in a group are greater in educational settings which emphasize being better than others and where evaluation standards are comparative (Anderman and Midgley 1997). It is, therefore, essential that teachers' beliefs and ensuing practices do not produce or reinforce competition and comparison among students but rather create a classroom atmosphere where students can experience success and satisfaction regardless of their ability level and feel a sense of personal accomplishment and growing mastery (see Bong and Skaalvik 2003). In addition to maintaining a collaborative classroom climate, negative classroom effects on academic self-concept may be lessened by well-developed teacher awareness of students' individual needs and by competence adapting tasks and activities to accommodate these needs in order to match the students' zone of proximal development (Pressley et al. 1996). Optimal differentiation and individuation is a challenging instructional task requiring both sensitivity to individual students' academic, motivational and emotional needs and at the same time, attention to the whole group of students and classroom dynamics. This kind of responsive orchestration at its best requires small enough class sizes. Learning methods which encourage students to build autonomy and agency and provide feedback concerning their individual improvement seem to be one method, particularly in mathematics, to reduce social comparison processes (see also Marsh and Craven 2002).

Limitations

This study has at least five limitations that should be taken into account. First, a clear limitation of the study was that in contrast to Grades 4 and 6 in which the self-concept scales of literacy and math consisted of several items, in Grade 3 the self-concept of ability scales consisted of one item, which is not psychometrically optimal. However, similar measures have been used also in previous studies focusing on early grades (e.g., Lazarides et al. 2018; Viljaranta et al. 2014). Furthermore, the relatively high correlations found both for self-concept in literacy

(across age correlations between .66 and .94) and for self-concept in math (across age/time correlations between .44 and .68) suggest that the measures succeeded in tapping self-concept comparably at the respective ages. With respect to Grade 3, the findings need to be replicated preferably by using a measure based on several items. Second, the math tests employed in the present study focused on mastery of basic arithmetic skills. Thus, the findings have to be replicated by using other math measures, such as those that focus on problem solving. Third, similar studies should be conducted also in countries with different education systems (e.g., in countries where tracking is more clearly present in schools and classrooms) to confirm whether the small evidence for BFLP effect that we found is due to the nonselective school system and not due to participants young age. Fourth, although we followed the same students from Grade 3 to Grade 6, it is important to emphasize that the results of the present study are based on correlational data collected at a single point in time. As such, we cannot suggest causal interpretations as a result of these models. Finally, all differences between girls and boys were not statistically significant according to the chi-square test, although the paths were typically significant only either for boys or girls. Therefore, studies with other samples are needed before generalizing the results concerning gender differences.

Our study extended previous research by examining BFLP effect already in early stage of school career. However, for fully understanding of BFLP effect, it would be important in future studies to examine also with different designs how students' self-concept develops in classroom context. For example, it would be important to examine in more detail whether students in one class compare themselves with all classmates when forming their self-concept or whether they compare themselves mainly with a smaller reference groups such as friends or classmates of same gender.

Conclusions

The present study extends previous findings on the big-fish-little-pond effect by showing that in a nonselective school system average classroom performance has only a relatively small negative association with boys' self-concept in mathematics already at an early phase of their school career. In literacy and among girls, the average classroom performance negative effect did not emerge until the end of primary school.

Acknowledgements Open access funding provided by University of Jyväskylä (JYU).

Funding This study was supported by grants from the Academy of Finland (Nos: 268 586, 265 817, 277 299, 276 239, 263 891, 299 506, 316 852, 323 773).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Anderman, E. M., & Midgley, C. (1997). Changes in achievement goal orientations, perceived academic competence, and grades across the transition to middle-level schools. *Contemporary Educational Psychology*, 22, 269–298. <https://doi.org/10.1006/ceps.1996.0926>.

- Atkinson, J. W. (1964). *An introduction to motivation*. Princeton, NJ: Van Nostrand.
- Aunola, K., & Räsänen, P. (2007). *The 3-minute basic arithmetic test*. Unpublished test material. Jyväskylä, Finland.
- Aunola, K., Leskinen, E., Onatsu-Arviolommi, T., & Nurmi, J.-E. (2002). Three methods for studying developmental change: a case of reading skills and self-concept. *British Journal of Educational Psychology*, *72*, 343–364. <https://doi.org/10.1348/000709902320634447>.
- Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York, NY: Freeman.
- Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: how different are they really? *Educational Psychology Review*, *15*, 1–40. <https://doi.org/10.1023/A:1021302408382>.
- Burnett, P. C. (2003). The impact of teacher feedback on student self-talk and self-concept in reading and mathematics. *Journal of Classroom Interaction*, *38*, 11–16.
- Calsyn, R. J., & Kenny, D. A. (1997). Self-concept of ability and perceived evaluation of others: cause or effect of academic achievement? *Journal of Educational Psychology*, *69*, 136–145. <https://doi.org/10.1037/0022-0663.69.2.136>.
- Chapman, J. W., & Tunmer, W. E. (1997). A longitudinal study of beginning reading achievement and reading self-concept. *British Journal of Educational Psychology*, *67*, 279–291. <https://doi.org/10.1111/j.2044-8279.1997.tb01244.x>
- Chiu, M. M., Chow, B. W.-Y., & Joh, S. W. (2017). Streaming, tracking and reading achievement: a multilevel analysis of students in 40 countries. *Journal of Educational Psychology*, *109*, 915–934. <https://doi.org/10.1037/edu0000188>.
- Cimpian, A. (2017). Early reasoning about competence is not irrationally optimistic, nor does it stem from inadequate cognitive representation. In A. J. Elliot, C. S. Dweck, & D. S. Yeager (Eds.), *Handbook of competence and motivation* (2nd ed., pp. 387–407). New York: Guilford.
- 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*, 430–447. <https://doi.org/10.1111/j.1467-8624.2007.01007.x>.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco, CA: Freeman.
- Eccles, J. S., Wigfield, A., Harold, R., & Blumenfeld, P. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development*, *64*, 830–847. <https://doi.org/10.2307/1131221>.
- Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon & R. M. Lerner (Series eds.) & N. Eisenberg (Volume ed.), *Handbook of child psychology*, Vol. 3, Social, emotional, and personality development (6th ed., pp. 1017–1095). New York, Wiley.
- Furtak, E. M., Kiemer, K., Circi, R. K., Swanson, R., de Leon, V., Morrison, D., & Heredia, S. C. (2016). Teachers' formative assessment abilities and their relationship to student learning: findings from a four-year intervention study. *Instructional Science*, *44*, 267–291. <https://doi.org/10.1007/s11251-016-9371-3>.
- Guo, J., Marsh, H. W., Parker, P. D., & Dicke, T. (2018). Cross-cultural generalizability of social and dimensional comparison effects on reading, math, and science self-concepts for primary school students using the combined PIRLS and TIMSS data. *Learning and Instruction*, *58*, 210–219. <https://doi.org/10.1016/j.learninstruc.2018.07.007>.
- Harter, S. (1982). The perceived competence scale for children. *Child Development*, *53*, 87–97. <https://doi.org/10.2307/1129640>.
- Holopainen, L., Ahonen, T., & Lyytinen, H. (2001). Predicting delay in reading achievement in a highly transparent language. *Journal of Learning Disabilities*, *34*, 401–413. <https://doi.org/10.1177/002221940103400502>.
- Hox, J. J., & Maas, C. J. M. (2002). Sample sizes for multilevel modeling. In J. Blasius, J. Hox, E. de Leeuw, & P. Schmidt (Eds.), *Social science methodology in the new millennium: proceedings of the fifth international conference on logic and methodology 2nd* (expanded ed., pp. 1–19). Opladen: RG: Leske+Budrich Verlag <http://www.geocities.com/joophox/publist/simnorm1.pdf>.
- 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*, 736–746. <https://doi.org/10.1016/j.lindif.2012.07.020>
- Koponen, T., & Räsänen, P. (2003). *NMART-Laskutaidon Testi [NMART counting skills test]*. Jyväskylä, Finland: Niilo Mäki Institute.
- Langdon, D. W., & Warrington, E. K. (1997). The abstraction of numerical relations: a role for the right hemisphere in arithmetic? *Journal of the International Neuropsychological Society*, *3*, 260–268.

- Lazarides, R., Viljaranta, J., Aunola, K., & Nurmi, J.-E. (2018). Teacher ability evaluation and changes in elementary student profiles of motivation and performance in mathematics. *Learning and Individual Differences*, *67*, 245–258. <https://doi.org/10.1016/j.lindif.2018.08.010>.
- Lerkanen, M.-K., Niemi, P., Poikkeus, A.-M., Poskiparta, E., Siekkinen, M., & Nurmi, J.-E. (2006–2016). *The First Step Study [Alkuportaati]*. Jyväskylä, Finland: University of Jyväskylä.
- Lindeman, J. (1998). *ALLU—Ala-asteen luketesti [ALLU—reading test for primary school]*. Turku, Finland: The Center for Learning Research, University of Turku.
- Lyytinen, H., Aro, M., Holopainen, L., Leiwo, M., Lyytinen, P., & Tolvanen, A. (2006). Children's language development and reading acquisition in a highly transparent orthography. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 47–62). Mahwah, NJ: Lawrence Erlbaum.
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, *79*, 280–295. <https://doi.org/10.1037/0022-0663.79.3.280>.
- Marsh, H. W., & Craven, R. G. (2002). The pivotal role of frames of reference in academic self-concept formation: the “big fish little pond” effect. In F. Pajares & T. Urdan (Eds.), *Adolescence and education* (pp. 83–123). Greenwich, CT: Information Age.
- 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*, 364–376. <https://doi.org/10.1037/0003-066X.58.5.364>.
- Marsh, H. W., & O'Mara, A. J. (2010). Long-term total negative effects of school-average ability on diverse educational outcomes. *Zeitschrift für Pädagogische Psychologie*, *24*, 51–72. <https://doi.org/10.1024/1010-0652/a000004>.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). The big-fish-little-pond effect: persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal*, *44*, 631–669. <https://doi.org/10.3102/0002831207306728>.
- Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J., & Graven, R. G. (2008). The big-fish-little-pond-effect stands up to critical scrutiny: implications for theory, methodology, and future research. *Educational Psychology Review*, *20*, 319–350. <https://doi.org/10.1007/s10648-008-9075-6>.
- Marsh, H. W., Abduljabbar, A. S., Morin, A. J. S., Parker, P., Abdelfattah, F., Nagengast, B., & Abu-Hilal, M. M. (2015). The big-fish-little-pond effect: generalizability of social comparison processes over two age cohorts from Western, Asian, and Middle Eastern Islamic countries. *Journal of Educational Psychology*, *107*, 258–271. <https://doi.org/10.1037/a0037485>.
- Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: academic self-concept, grades, test scores, and tracking over 6 years. *Developmental Psychology*, *54*, 263–280. <https://doi.org/10.1037/dev0000393>.
- Muenks, K., Wigfield, A., & Eccles, J. S. (2018). I can do this! The development and calibration of children's expectations for success and competence beliefs. *Developmental Review*, *48*, 24–39. <https://doi.org/10.1016/j.dr.2018.04.001>.
- Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
- National Core Curriculum. (2014). Finnish National Board of Education National Core Curriculum. Retrieved from http://oph.fi/english/curricula_and_qualifications/basic_education
- Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child Development*, *49*, 800–814. <https://doi.org/10.2307/1128250>.
- Nurmi, J.-E., & Aunola, K. (2005). Task-motivation during the first school years. A person-oriented approach to longitudinal data. *Learning and Instruction*, *15*, 103–122. <https://doi.org/10.1016/j.learninstruc.2005.04.009>.
- Parker, P. D., Marsh, H. W., Lüdtke, O., & Trautwein, U. (2013). Differential school contextual effects for mathematics and English: integrating the big-fish-little-pond effect and the internal/external frame of reference. *Learning and Instruction*, *23*, 78–89. <https://doi.org/10.1016/j.learninstruc.2012.07.001>.
- Pesu, L., Viljaranta, J., & Aunola, K. (2016). The role of parents' and teachers' beliefs in children's self-concept development. *Journal of Applied Developmental Psychology*, *44*, 63–71. <https://doi.org/10.1016/j.appdev.2016.03.001>.
- Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: theory, research and application* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Pinxten, M., Wouters, S., Preckel, F., Niepel, C., De Fraine, B., & Verschueren, K. (2015). The formation of academic self-concept in elementary education: a unifying model for external and internal comparisons. *Contemporary Educational Psychology*, *41*, 124–132. <https://doi.org/10.1016/j.cedpsych.2014.12.003>.
- Plieninger, H., & Dickhäuser, O. (2015). The female fish is more responsive: gender moderates the BFLPE in the domain of science. *Educational Psychology*, *35*, 213–227. <https://doi.org/10.1080/01443410.2013.814197>.

- Preckel, F., Zeidner, M., Goetz, T., & Schleyer, E. J. (2008). Female “big fish” swimming against the tide: the “big-fish-little-pond effect” and gender-ratio in special gifted classes. *Contemporary Educational Psychology*, *33*, 78–96. <https://doi.org/10.1016/j.cedpsych.2006.08.001>.
- Pressley, M., Hogan, K., Wharton-McDonald, R., Mistretta, J., & Ettenberg, S. (1996). The challenges of instructional scaffolding: the challenges of instruction that supports student thinking. *Learning Disabilities Research & Practice*, *11*, 38–146.
- Räsänen, P., Salminen, J., Wilson, A., Aunio, P., & Dehaene, S. (2009). Computer-assisted intervention for children with low numeracy skills. *Cognitive Development*, *24*, 450–472. <https://doi.org/10.1016/j.cogdev.2009.09.003>.
- Shavelson, R., Hubner, J., & Stanton, G. C. (1976). Self-concept: validation of construct interpretations. *Review of Educational Research*, *46*, 407–441. <https://doi.org/10.2307/1170010>.
- Shurtz, 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>.
- Statistics Finland (2007). *Statistical database*. Retrieved from http://www.stat.fi/tup/tilastotietokannat/index_en.html
- Thijs, J., Verkuyten, M., & Helmond, P. (2010). A further examination of the big-fish-little-pond effect: perceived position in class, class size, and gender comparisons. *Sociology of Education*, *83*, 333–345. <https://doi.org/10.1177/0038040710383521>.
- Torppa, M., Lyytinen, P., Erskine, J., Eklund, K., & Lyytinen, H. (2010). Language development, literacy skills, and predictive connections to reading in Finnish children with and without familial risk for dyslexia. *Journal of Learning Disabilities*, *43*, 308–321. <https://doi.org/10.1177/0022219410369096>.
- 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*, 788–806. <https://doi.org/10.1037/0022-0663.98.4.788>.
- Trautwein, U., Gerlach, E., & Lüdtke, O. (2008). Athletic classmates, physical self-concept, and free-time physical activity: a longitudinal study of frame of reference effects. *Journal of Educational Psychology*, *100*, 988–1001. <https://doi.org/10.1037/0022-0663.100.4.988>.
- Viljaranta, J., Tolvanen, A., Aunola, K., & Nurmi, J.-E. (2014). The developmental dynamics between interest, self-concept of ability, and academic performance. *Scandinavian Journal of Educational Research*, *58*, 734–756. <https://doi.org/10.1080/00313831.2014.904419>.
- Wigfield, A., Eccles, J. S., Schiefele, U., Roeser, R. W., & Davis-Kean, P. (2006). Development of achievement motivation. In W. Damon & R. M. Lerner (Series eds.) & N. Eisenberg (Volume ed.), *Handbook of child psychology*, Vol. 3, Social, emotional, and personality development (6th ed., pp. 933–1002). New York, Wiley.
- Wouters, S., Colpin, H., Van Damme, J., De Laet, S., & Verschueren, K. (2013). Early adolescents' academic self-concept formation: do classmates or friends matter most? *Learning and Individual Differences*, *27*, 193–200. <https://doi.org/10.1016/j.lindif.2013.09.002>.
- Zhang, X., Koponen, T., Räsänen, P., Aunola, K., Lerkkanen, M., & Nurmi, J. (2014). Linguistic and spatial skills predict early arithmetic development via counting sequence knowledge. *Child Development*, *85*, 1091–1107. <https://doi.org/10.1111/cdev.12173>.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Kati Vasalampi. Department of Psychology, University of Jyväskylä, P. O. Box 35, 40014 University of Jyväskylä, Finland. E-mail: kati.vasalampi@jyu.fi

Current themes of research:

Learning motivation, student engagement, school well-being, school dropout.

Most relevant publications in the field of Psychology of Education:

- Pöysä, S., Vasalampi, K., Muotka, J., Lerkkanen, M.-K., Poikkeus, A.-M., & Nurmi, J.-E. (2019). *Teacher-student interaction and lower secondary school students' situational engagement*. *British Journal of Educational Psychology*, 89, 374-392.
- Torppa, M., Vasalampi, K., Eklund, K., Sulkunen, S., Niemi, P., & Ahonen, T. (2019). *Reading comprehension difficulty is often distinct from difficulty in reading fluency and accompanied with problems in motivation and school well-being*. *Educational Psychology*.
- Virtanen, T., Vasalampi, K., Torppa, M., Lerkkanen, M.-K., & Nurmi, J.-E. (2019). *Changes in students' psychological well-being during transition from primary school to lower secondary school: A person-centered approach*. *Learning and Individual Differences*, 69, 138-149.
- Vasalampi, K., Kiuru, N., & Salmela-Aro, K. (2018). *The role of a supportive interpersonal environment and education-related goal motivation during the transition beyond upper secondary education*. *Contemporary Educational Psychology*, 55, 110-119.
- Vasalampi, K., Muotka, J., Pöysä, S., Lerkkanen, M.-K., Poikkeus, A.-M. & Nurmi, J.-E. (2016). *Assessment of students' situation-specific classroom engagement by an InSitu Instrument*. *Learning and Individual Differences*, 52, 46-52.

Eija Pakarinen. Department of Teacher Education, University of Jyväskylä, P. O. Box 35, 40014 University of Jyväskylä, Finland. E-mail: eija.k.pakarinen@jyu.fi

Current themes of research:

Teacher-student interactions, teacher-student relationship, teaching practices, teacher stress, motivation.

Most relevant publications in the field of Psychology of Education:

- Pakarinen, E., & Kikas, E. (2019). *Child-centered and teacher-directed practices in relation to calculation and word problem solving skills*. *Learning and Individual Differences*, 70, 76–85.
- Engels, M., Pakarinen, E., Lerkkanen, M.-K., & Verschueren, K. (2019). *Students' academic and emotional adjustment during the transition from primary to secondary school: a cross-lagged study*. *Journal of School Psychology*, 76, 140–158.
- Muhonen, H., Pakarinen, E., Poikkeus, A.-M., Rasku-Puttonen, H., & Lerkkanen, M.-K. (2018). *Quality of educational dialogue and association with students' academic performance*. *Learning and Instruction*, 55, 67-79.
- Pakarinen, E., Silinskas, G., Hamre, B. K., Metsäpelto, R.-L., Lerkkanen, M.-K., Poikkeus, A.-M., & Nurmi, J.-E. (2018). *The cross-lagged associations between problem behaviors and teacher-student relationships in early adolescence*. *The Journal of Early Adolescence*, 38, 1100–1141.
- Pakarinen, E., Aunola, K., Kiuru, N., Lerkkanen, M.-K., Poikkeus, A.-M., Siekkinen, M., & Nurmi, J.-E. (2014). *The cross-lagged associations between classroom interactions and children's achievement behaviors*. *Contemporary Educational Psychology*, 39, 248–261.

Minna Torppa. Department of Teacher Education, University of Jyväskylä, P. O. Box 35, 40014 University of Jyväskylä, Finland. E-mail: minna.p.torppa@jyu.fi

Current themes of research:

Reading and spelling development, school well-being, learning difficulties.

Most relevant publications in the field of Psychology of Education:

- Torppa, M., Niemi, P., Vasalampi, K., Poikkeus, A.-M., & Lerkkanen, M.-K. (2019). *Leisure reading (but not any kind) and reading comprehension support each other – A longitudinal study across Grades 1 and 9*. *Child Development*.
- Georgiou, G., Torppa, M., Landerl, K., Manolitsis, G., de Jong, P., & Parrila, R. (2019). *Reading and spelling development across languages varying in orthographic consistency: Do their paths cross?* *Child Development*.

- Torppa, M., Eklund, K., Sulkunen, S., Niemi, P., & Ahonen, T. (2018). *Why do boys and girls perform differently on PISA Reading in Finland? The effects of reading fluency, achievement behaviour, leisure reading and homework activity*. *Journal of Research in Reading*, 41, 122-139.
- Parhiala, P., Torppa, M., Vasalampi, K., Eklund, K., Poikkeus, A.-M., & Aro, T. (2018). *Profiles of school motivation and emotional well-being among adolescents: Associations with math and reading performance*. *Learning and Individual Differences*, 61, 196-204.
- Koponen, T., Salmi, P., Torppa, M., Lerkkanen, M.-K., Poikkeus, A.-M., & Nurmi, J.-E. (2016). *Counting and rapid naming predict the fluency of arithmetic and reading skills*. *Contemporary Educational Psychology*, 44-45, 83-94.

Jaana Viljaranta. School of Educational Sciences and Psychology, Philosophical Faculty, University of Eastern Finland, P. O. Box 111, 80101 Joensuu. E-mail: jaana.viljaranta@uef.fi

Current themes of research:

The role of child-teacher and child-parent interaction in development of learning motivation and skills, and the role that child's individual characteristics, such as temperament, play in this interaction.

Most relevant publications in the field of Psychology of Education:

- Dietrich, J., Moeller, J., Guo, J., Viljaranta, J., & Kracke, B. (2019). *In-the-moment profiles of expectancies, task values, and costs*. *Frontiers in Psychology*.
- Viljaranta, J., Silinskas, G., Lerkkanen, M.-K., Hirvonen, R., Pakarinen, E., Poikkeus, A.-M., & Nurmi, J.-E. (2018). *Maternal homework assistance and children's task-persistent behavior in elementary school*. *Learning and Instruction*.
- Dietrich, J., Viljaranta, J., Moeller, J., & Kracke, B. (2017). *Situational expectancies and task values as predictors of students' effort*. *Learning and Instruction*, 47, 53-64.
- Lazarides, R., Viljaranta, J., Ranta, M., & Salmela-Aro, K., (2017). *Adolescents' preparedness and motivation across the transition to post-comprehensive education*. *Contemporary Educational Psychology*.
- Viljaranta, J., Aunola, K., Mullola, S., Virkkala, J., Hirvonen, R., Pakarinen, E., & Nurmi, J.-E. (2015). *The role of temperament on children's skill development: Teachers' Interaction Styles as Mediators*. *Child Development*, 86, 1191-1209.

Marja-Kristiina Lerkkanen. Department of Teacher Education, University of Jyväskylä, P. O. Box 35, 40014 University of Jyväskylä, Finland. E-mail: marja-kristiina.lerkkanen@jyu.fi

Current themes of research:

The quality of teacher-child interaction in learning and motivation, reading skills development, motivation for learning, teaching practices.

Most relevant publications in the field of Psychology of Education:

- Lerkkanen, M.-K. (2018). *The influence of instruction on reading motivation in Finland*. In P. Orellana Garcia and P. Baldwin Lind & (Eds.). *Reading achievement and motivation in boys and girls: field studies and methodological approaches*. *Literacy Studies*, 15. Springer Education.
- Lerkkanen, M.-K., Kiuru, N., Pakarinen, E., Poikkeus, A.-M., Rasku-Puttonen, H., Siekkinen, M., & Nurmi, J.-E. (2016). *Child-Centered versus Teacher-Directed Teaching Practices: Associations with the Development of Academic Skills in the First Grade at School*. *Early Childhood Research Quarterly*, 36(3), 145-156.
- Muhonen, H., Pakarinen, E., Poikkeus, A.-M., Rasku-Puttonen, H., & Lerkkanen, M.-K. (2018). *Quality of educational dialogue and association with students' academic performance*. *Learning and Instruction*, 55, 67-79.
- Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Salminen, J., Silinskas, G., Siekkinen, M., & Nurmi, J.-E. (2017). *Longitudinal associations between teacher-child interactions and academic skills in elementary school*. *Journal of Applied Developmental Psychology*, 52, 191-202.

Soodla, P., Lerkkanen, M.-K., Niemi, P., Kikas, E., Silinskas, G., & Nurmi, J.-E. (2015). *Does early reading instruction promote the rate of acquisition? A comparison of two transparent orthographies*. *Learning and Instruction*, 38, 14-23.

Anna-Maija Poikkeus. Department of Teacher Education, University of Jyväskylä, P. O. Box 35, 40014 University of Jyväskylä, Finland. E-mail: anna-maija.poikkeus@jyu.fi

Current themes of research:

Trajectories of young children's reading and math achievement and motivation, development of language and academic skills of children with reading disability, teacher-student relationships, classroom interaction and student engagement, dialogical teaching and development of agency, social skills and peer relations.

Most relevant publications in the field of Psychology of Education:

- Jääskelä, P., Poikkeus, A.-M., Vasalampi, K., Valleala, U. M. & Rasku-Puttonen, H. (2016) *Assessing agency of university students. Validation of the AUS Scale*. *Studies in Higher Education*, 42, 2061-2079.
- Kiuru, N., Aunola, K., Lerkkanen, M.-K., Pakarinen, E., Poskiparta, E., Ahonen, T., Poikkeus, A.-M., & Nurmi, J.-E. (2015). *Positive teacher and peer relations combine to predict primary school students' academic skill development*. *Developmental Psychology*, 51, 434-446.
- Lyytinen, H., Erskine, J., Tolvanen, A., Torppa, M., Poikkeus, A.-M., & Lyytinen, P. (2006). *Trajectories of reading development. A follow-up from birth to school age of children with and without risk for dyslexia*. *Merrill-Palmer Quarterly*, 52, 514-546.
- Metsäpelto, R.-L., Pakarinen, E., Kiuru, N., Poikkeus, A.-M., Lerkkanen, M.-K., & Nurmi, J.-E. (2015). *Developmental Dynamics Between Children's Externalizing Problems, Task-Avoidant Behavior, and Academic Performance in Early School Years: A 4-Year Follow-Up*. *Journal of Educational Psychology*, 107 (1), 246-257.
- Muhonen, H., Pakarinen, E., Poikkeus, A.-M., Lerkkanen, M.-K., & Rasku-Puttonen, H. (2018). *Quality of educational dialogue and association with students' academic performance*. *Learning and Instruction*, 55, 67-79.