



# Changes in students' self-efficacy when learning a new topic in mathematics: a micro-longitudinal study

Karin E. S. Street<sup>1</sup> · Lars-Erik Malmberg<sup>2</sup> · Gabriel J. Stylianides<sup>2</sup>

Accepted: 30 May 2022 / Published online: 14 July 2022  
© The Author(s) 2022

## Abstract

Self-efficacy in mathematics is related to engagement, persistence, and academic performance. Prior research focused mostly on examining changes to students' self-efficacy across large time intervals (months or years), and paid less attention to changes at the level of lesson sequences. Knowledge of how self-efficacy changes during a sequence of lessons is important as it can help teachers better support students' self-efficacy in their everyday work. In this paper, we expanded previous studies by investigating changes in students' self-efficacy across a sequence of 3–4 lessons when students were learning a new topic in mathematics ( $n_{\text{Students}} = 170$ ,  $n_{\text{Time-points}} = 596$ ). Nine classes of Norwegian grade 6 ( $n = 77$ ) and grade 10 students ( $n = 93$ ) reported their self-efficacy for easy, medium difficulty, and hard tasks. Using multilevel models for change, we found (a) change of students' self-efficacy across lesson sequences, (b) differences in the starting point and change of students' self-efficacy according to perceived task difficulty and grade, (c) more individual variation of self-efficacy starting point and change in association with harder tasks, and (d) students in classes who were taught a new topic in geometry had stronger self-efficacy at the beginning of the first lesson as compared to those who were taught a new topic in algebra (grade 10), and students in classes who were taught a new topic in fractions had steeper growth across the lesson sequence as compared to those who were taught a new topic in measurement (grade 6). Implications for both research and practice on how new mathematics topics are introduced to students are discussed.

**Keywords** Self-efficacy · Mathematics · Change · Levels of task difficulty · Mathematics topics · Multilevel models for change

## 1 Introduction

*Self-efficacy* is broadly defined as the belief individuals hold regarding their own capability to carry out future tasks (Bandura, 1977). Domain-specific self-efficacy in mathematics is important because it predicts students' adaptive learning behaviors such as engagement,

---

✉ Karin E. S. Street  
karin.street@hvl.no

<sup>1</sup> Western Norway University of Applied Sciences, Sogndal, Norway

<sup>2</sup> University of Oxford, Oxford, UK

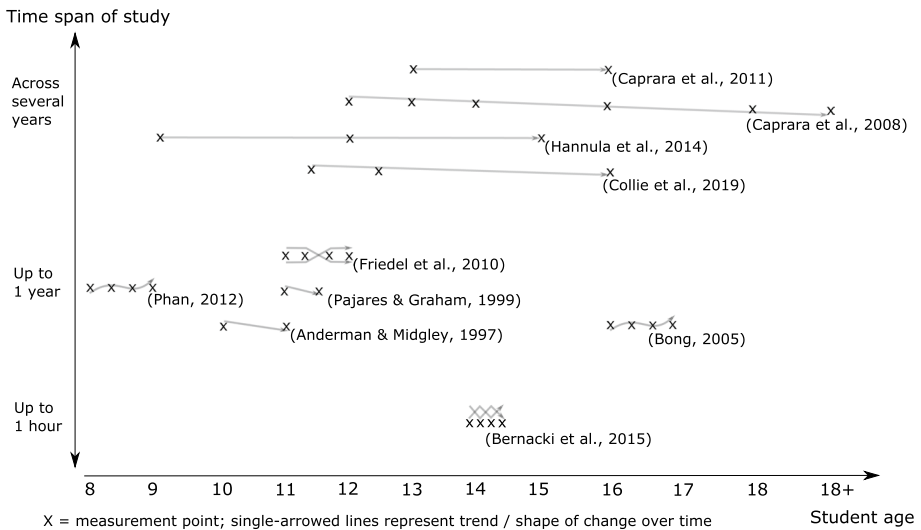
effort, enjoyment of mathematics, academic performance in mathematics, and future engagement with mathematics through study and career choices (e.g., Bandura, 1997; Hackett & Betz, 1989; Klassen, 2004; Klassen & Usher, 2010; Pajares, 1996). Given this importance of students' self-efficacy in their mathematical education, understanding how self-efficacy changes over time and what factors influence the strength of self-efficacy is crucial and can have implications for teaching practice. This paper aims to contribute to research knowledge in this area by investigating changes in students' self-efficacy across a sequence of lessons when students were learning a new topic in mathematics.

Self-efficacy has been conceptualized as a construct which can vary across three dimensions: *facet-specificity* (i.e., generality or specificity of the task at hand), *level of difficulty* (i.e., whether the task is perceived as easy or hard), and *strength* (i.e., students' degree of confidence in their capability to perform the task) (Bandura, 1997; Street et al., 2017). Students consider the specific task or challenge at hand (facet), including how difficult they perceive the task to be (level), when they formulate the strength of their self-efficacy beliefs (see Street et al., 2017). While there are some studies on level-specific mathematics self-efficacy (e.g., Chen & Zimmerman, 2007; Liu et al., 2020; Street et al., 2017; Zimmerman & Martinez-Pons, 1990a, b), to the best of our knowledge, there are no studies that investigated *changes* in students' level-specific mathematics self-efficacy.

In this paper, we investigated changes in mathematics self-efficacy according to the level of task difficulty across sequences of lessons in which a new topic was introduced. This focus is important because it can help us understand how to best support adaptive student self-efficacy in a challenging learning situation for both students and teachers in the classroom, such as when students are learning a new mathematics topic. Our study expands previous research on students' mathematics-specific self-efficacy in at least four ways.

First, we investigated self-efficacy changes during a sequence of lessons in typical classroom settings, whereas prior research, notably Bernacki et al. (2015), investigated short-term changes in mathematics self-efficacy during a sequence of computerized learning tasks. Second, while longer-term longitudinal investigations of students' self-efficacy exist (e.g., Bong, 2005; Collie et al., 2019; see Fig. 1), these studies included measurements of self-efficacy that ranged from six months to several years apart, thereby providing little insight into the process of self-efficacy change which can happen in shorter time intervals. Without micro-longitudinal studies, we cannot know if the changes in self-efficacy in the aforementioned studies happened gradually and in a unidirectional way or otherwise. Third, while previous studies have typically focused on efficacy beliefs about mathematics in general, we investigated self-efficacy beliefs for tasks of different levels of difficulty. Fourth, while previous studies confirmed that students' self-efficacy differ as a function of domain and task specificity (e.g., Bong, 2001; Bong & Hocevar, 2002; Hornstra et al., 2016), no studies that we know of investigated the role of mathematical content areas (algebra, geometry, etc.) in relation to changes in students' self-efficacy. In our study, we explored the role of different content areas on students' self-efficacy at the starting point<sup>1</sup> and over a sequence of lessons, in order to see whether content area is a factor that future research needs to consider more systematically when investigating changes in students' mathematics self-efficacy over time.

<sup>1</sup> In the literature on growth modelling the terms "level" and "change" are typically used for interpreting the statistical terms "intercept" and "slope". As we use the term "level" to refer to level of difficulty, we will use the term "starting-point" for interpreting the intercept, as it denotes the beginning of a sequence of lessons.



**Fig. 1** Overview of empirical findings on self-efficacy change

## 2 Development of mathematics self-efficacy

Students' appraisal of their mastery experiences is considered the strongest of the four proposed sources of self-efficacy (Byars-Winston et al., 2017; Usher & Pajares, 2008). Furthermore, there is a theoretically proposed mutual relationship between mathematics self-efficacy and performance, where students' self-efficacy influences their performance, which in turn influences their self-efficacy (Bandura, 1997; Talsma et al., 2018). Importantly, academic performance does not influence self-efficacy directly, but is cognitively appraised by the individual, i.e., an experience of learning or mastery. Students make sense of their performance experiences post hoc, in a process which likely includes other belief structures, such as students' beliefs about the nature of ability (i.e., whether ability is innate or can be developed; Dweck, 2002). Mastery experiences from tasks of high perceived similarity provide the strongest source for students' concurrent self-efficacy (e.g., Byars-Winston et al., 2017), and changes in self-efficacy are most likely to occur when students learn new things or engage with new experiences (Usher & Pajares, 2008). Borgonovi and Pokropek (2019), using the PISA 2012 data ( $N=290,738$ ), found an effect of task exposure on 15 to 16-year-old students' self-efficacy, where students reported higher self-efficacy in association with more familiar tasks. Thus, when new topics are introduced during classroom learning, these might give rise to uncertainty and fluctuations in students' mathematics self-efficacy.

### 2.1 Time span of self-efficacy change in previous studies

Change can be studied across longer or shorter periods of time. A longer-term perspective (e.g., measurement points across months or years) can provide insight into self-efficacy changes in relation to factors such as student cognitive development or school and teacher

changes, while a micro-analytic focus (e.g., measurement points across learning events) can provide insight into the processes through which self-efficacy develops during and across classroom learning in mathematics. In Fig. 1, we provide a schematic overview of empirical findings from the studies we reviewed, including time frame (on *y*-axis), number of measurement points (indicated as “*x*” along the *x*-axis for each study), and trend/shape of change reported (indicated by single-headed lines for each study). We found only one study that investigated changes in mathematics self-efficacy across several years, at ages 11, 12, and 16 (Collie et al., 2019). Due to the scarcity of such studies, we included three longitudinal studies that investigated related concepts across several years, for instance from age 13 to 16 (academic self-efficacy; Caprara et al., 2011), across ages 12 to 22 (self-efficacy for self-regulation; Caprara et al., 2008), or from age 9, to 12, to 15 (mathematics self-concept; Hannula et al., 2014). Overall, these studies indicated significant rank order stability (the degree to which students’ order in a class, as compared to their peers, is stable across time; coefficients ranging from 0.30 to 0.49) but also trends of decline in terms of mean strength (growth slopes ranging from  $-0.02$  to  $-0.04$  per year) of students’ beliefs over time. Consistent with these trends, several studies have included measures of students’ mathematics self-efficacy at different ages, in a correlational design, and reported weaker self-efficacy for older students (see, e.g., Bong, 2009; Wigfield et al., 1998; Zimmerman & Martinez-Pons, 1990a, b).

Most commonly, changes in mathematics self-efficacy have been investigated over the course of a year (e.g., Anderman & Midgley, 1997; Bong, 2005; Friedel et al., 2010; Pajares & Graham, 1999; Phan, 2012), with the time between measurements varying from six months (e.g., Friedel et al., 2010) to a year (Anderman & Midgley, 1997). Findings from these studies varied, including negative change (e.g., Anderman & Midgley, 1997; Pajares & Graham, 1999), no overall change (e.g., Friedel et al., 2010), positive growth (e.g., Bong, 2005; Phan, 2012), and nonlinear change trajectories (e.g., Bong, 2005; Phan, 2012) in students’ self-efficacy. Furthermore, Bong’s (2005) study, which extended over one year, reported significant rank order stability of students’ mathematics self-efficacy (coefficients ranging from 0.29 to 0.47), which is comparable to the stability coefficients reported by the studies conducted over several years.

Despite the variations in findings, the studies we reviewed demonstrated overall trends of fluctuations, where students’ mathematics self-efficacy fluctuated according to important external events, e.g., around exam times (Bong, 2005; Pajares & Graham, 1999) or across school changes (Anderman & Midgley, 1997; Friedel et al., 2010). Thus, findings from these studies indicated a dynamic nature of students’ mathematics self-efficacy, increasingly evident with more frequent measurement occasions. However, while some of these previous studies investigated more intensively changes in students’ self-efficacy, the studies were not sensitive enough to capture fluctuations in students’ self-efficacy as they engaged in daily learning tasks in the classroom. Furthermore, when self-efficacy examinations are connected to particular events such as exams or change of school/teacher, they give little insight into fluctuations of self-efficacy during normal teaching practice, which is important in terms of informing classroom practice more broadly.

## 2.2 Self-efficacy change across tasks

We found only one study (Bernacki et al., 2015) that investigated changes in self-efficacy across learning tasks. Bernacki et al. (2015) investigated US 9th grade students’ ( $N=107$ ) task-specific self-efficacy and problem solving performance while working on algebra

problems within an intelligent tutoring system (“Cognitive Tutor Algebra”). A single self-efficacy item (*How confident are you that you could solve a math question like this one in the future?*) was used to measure students' self-efficacy after the completion of every fourth algebra problem in the unit. A total of four self-efficacy prompts were given across the learning unit, where the students typically responded to a self-efficacy prompt every five minutes. The authors reported that self-efficacy was quite stable across time (correlation estimates ranging from  $r=0.61$  to  $r=0.86$ ). Furthermore, at the group level, self-efficacy did not change between the time points. However, the authors found that, within a single learning task, self-efficacy “varied reliably” at least once for 60% of the students. This indicated that some students experienced increases in self-efficacy and other students experienced decreases in self-efficacy, resulting in no overall mean change. Bernacki et al.'s (2015) study indicated that there is significant variation in students' self-efficacy across problem-solving tasks, even when group-level changes are not evident. While Bernacki et al. (2015) provided a micro-analytic approach during students' learning of mathematics, their study focused on learning within an intelligent tutoring system, which is not typical in everyday mathematics practice. Furthermore, Bernacki et al. (2015) used a single self-efficacy item, which means it is hard to disattenuate measurement error from “actual” change in self-efficacy over time.

### 2.3 The role of perceived task difficulty

While some previous studies reported that the strongest association between self-efficacy and performance outcomes was with medium difficulty or hard tasks (Marshall & Brown, 2004; Street et al., 2017), a recent study reported declining correlations between Chinese eighth-grade students' mathematics self-efficacy and subsequent problem posing performance with increasing task difficulty (Liu et al., 2020). Liu et al. (2020) argued that students might form their self-efficacy on the basis of how easy or hard the task appears, rather than through assessing their own performance capability. This is likely to be especially relevant when students learn something new. In terms of the role of familiarity and task difficulty on the formation of self-efficacy, Borgonovi and Pokropek (2019) found an association between task difficulty (as indicated by the mean of participants' self-efficacy beliefs) and to what degree student mathematics achievement moderated the effect of previous task exposure on students' self-efficacy. Exposure on harder tasks had a stronger effect on self-efficacy for lower achieving students, while exposure on easier tasks had a stronger effect on the self-efficacy for higher achieving students. As these previous studies were correlational, we can say little about lesson-by-lesson changes in mathematics self-efficacy for different levels of difficulty.

As argued by Street et al. (2017), it is important to determine whether changes in students' self-efficacy differ as a function of levels of perceived difficulty. Also, as we mentioned earlier, cognitively appraised performance experiences are the strongest source of students' mathematics self-efficacy (Byars-Winston et al., 2017). Students' beliefs about the relationship between task difficulty, effort, and ability (Nicholls, 1978) are likely to play a role in these performance appraisals. For instance, succeeding on hard tasks might make a stronger impact on students' self-efficacy relative to succeeding on easy tasks, as performance on hard tasks is likely to be seen as more informative in terms of future performance capability. At the same time, tasks of high perceived difficulty present a higher risk of failure, and might thus also deter students from engaging with them in the first place. Another important appraisal factor might be the memorability of the performance

event: To what degree a lesson experience “stands out” might be related to whether and how it has an effect on the formation of students’ future self-efficacy (Butz & Usher, 2015; Stylianides & Stylianides, 2014). In this view, student struggle on relatively harder tasks (see, e.g., Bobis et al., 2021; Warshauer, 2015) might provide a memorable mastery experience of high importance. Investigating the role of task difficulty on students’ mathematics self-efficacy starting point (at the beginning of a lesson sequence) and change (across the lesson sequence) is thus important because it can provide information regarding potential effects of task selection in the classroom. For instance, it can help inform decisions of whether teachers might be better off directing students towards easier tasks (in order to maximize the chance for success), or to focus on harder tasks (to maximize the memorability and importance of the experience).

## 2.4 The role of mathematical content areas

Dawkins and Karunakaran (2016) drew attention to the role of mathematical content area (e.g., algebra, geometry, analysis), including students’ mathematical meanings for mathematical content, as a potentially influential factor of students’ proof-related behavior. Similarly, Kücheman and Hoyles (Küchemann & Hoyles, 2001) found that students’ performance in mathematical reasoning is not uniform across content areas and that students’ responses to unfamiliar geometrical reasoning items were subject to more variation between classes, as compared to more familiar algebraic reasoning items. In terms of self-efficacy, Bong (1997) reported that differences between students’ self-efficacy in algebra (arithmetic progression problems) and physics (constant acceleration problems) were related to perceived differences/similarities between the two classes of problem. According to Bong (1997), similarity perceptions might be part of the mechanism through which self-efficacy beliefs are generalized. In a more recent study, Krawitz and Schukajlow (2018) found differences in students’ mathematics self-efficacy according to problem type (i.e., lower self-efficacy for modelling versus dressed up word or intra-mathematical problems) but not topic (Pythagorean theorem versus linear functions).

Norwegian schools follow the national curriculum in mathematics, outlining main content areas and learning goals for different grade levels. The LK-06 national curriculum (The Norwegian Directorate for Education and Training, 2006) (which was relevant for Norwegian grade 6 and grade 10 students for 8 years prior to 2014, the year when the data for our study were collected) indicated at which stage students were first introduced to different content areas and main topics within those areas. We investigated students’ self-efficacy when they were exposed to a new mathematical topic and this topic could be in a number of different content areas. Thus, although the topic would be new, students’ familiarity with different aspects of the topic might differ depending on their previous exposure with the particular mathematical content area as per the mathematics curriculum they go through, and this familiarity could potentially make a difference in their self-efficacy. Considering previous mastery experiences are key to the development of students’ mathematics self-efficacy (Byars-Winston et al., 2017), it is possible, the characteristics as well as the familiarity of the content of the mathematics lessons are related to both the starting point and the change of students’ self-efficacy during a lesson sequence. Despite previous studies asserting the domain and task specificity of students’ self-efficacy beliefs (see, e.g., Bong, 1997, 2001; Bong & Hocevar, 2002; Hornstra et al., 2016; Liu et al., 2020), to our knowledge, no studies investigated the role of mathematical content areas in relation to changes in self-efficacy over time.

It would be important to ascertain whether there are differences in the changes of students' mathematics self-efficacy as a function of content area, especially in terms of the need for targeted teacher support. Furthermore, for longitudinal studies, it is important to consider whether changes in self-efficacy as a function of, for instance, classroom learning or age could be conflated with changes in self-efficacy due to changes in mathematics content.

## 2.5 Summary

Despite the theoretical idea that appraisals of mastery experiences from key learning events are central to the formation of and change in self-efficacy (Bandura, 1997), we found only one study (Bernacki et al., 2015) that investigated self-efficacy change with a micro-longitudinal design, and no studies that investigated such change in an ecological setting (during regular classroom learning), nor in relation to perceived levels of difficulty of the tasks. Investigating the role of perceived difficulty on students' mathematics self-efficacy at the starting point and across a lesson sequence is important because it could have implications for the selection of tasks in the classroom. To help address gaps in existing knowledge, we aimed in the current study to investigate how students' self-efficacy changes across regular classroom lessons in mathematics when students are introduced to a new topic, and whether the variability of self-efficacy differs by levels of perceived difficulty. Furthermore, we aimed to explore the role of mathematics content area on the starting point and change of students' self-efficacy. To enable a micro-analytic investigation into the possible change of mathematics self-efficacy across a sequence of lessons, we collected measures of multidimensional self-efficacy from students as they were introduced to a new topic in mathematics.

## 3 Research questions

Below we present the research questions for our study, which follow from the previous discussion. We also present our hypotheses for each question as relevant.

- (1) What is the starting point and change of students' mathematics self-efficacy for easy, medium difficulty, and hard tasks during a lesson sequence of a new topic?

We expected lower self-efficacy strength for harder, as compared with easier, perceived tasks at the beginning of lesson one (Marshall & Brown, 2004; Street et al., 2017) (Hypothesis 1). While we expected to see fluctuations in mean self-efficacy strength across the lessons (Bernacki et al., 2015; Bong, 2005) (Hypothesis 2), we did not pose any hypotheses regarding the shape or magnitude of these fluctuations, in relation to levels of perceived difficulty. The reason for not posing a specific hypothesis is the lack of previous empirical studies at the micro-longitudinal level of analysis and that we did not find it appropriate to extrapolate findings from longer-term studies to the timeframe of our study. (The same logic applies also for aspects of research questions 2, 3, and 4 for which we did not pose any hypotheses.)



- (2) Are there individual differences in the starting point and change of students' mathematics self-efficacy for easy, medium difficulty, and hard tasks during a lesson sequence of a new topic?

Building on Phan (2012) and Bernacki (2015), we predicted credible variance in starting point and change of students' mathematics self-efficacy (Hypothesis 3). We did not pose any hypotheses regarding individual differences in relation to perceived level of difficulty.

- (3) What is the starting point and change of students' mathematics self-efficacy for easy, medium difficulty, and hard tasks, according to student grade?

We expected older students to experience lower self-efficacy at the beginning of the first lesson as compared with younger students (Bong, 2009; Wigfield et al., 1998; Zimmerman & Martinez-Pons, 1990a, 1990b) (Hypothesis 4). We did not pose any hypotheses regarding the shape or magnitude of students' self-efficacy change in relation to student grade.

- (4) What is the starting point and change of students' mathematics self-efficacy for easy, medium difficulty, and hard tasks, according to mathematics content areas?

Despite a lack of previous empirical investigations in this area, we expected content area characteristics and familiarity could have an effect on the starting point and change of students' mathematics self-efficacy, similarly to how mathematical content was found to be a potentially influential factor of students' proof-related behavior (e.g., Dawkins & Karunakaran, 2016; Kücheman & Hoyles, 2001). We did not pose any hypotheses regarding the effect of mathematical content or students' perceived level of difficulty on students' changes of self-efficacy.

## 4 Method

### 4.1 Design

In order to investigate change over time in students' mathematics self-efficacy over the course of a sequence of lessons in mathematics (3–4 lessons, lasting 1–2 weeks), we applied a micro-analytic longitudinal design, including students in four primary and five secondary school classes in Norway ( $N=170$ ). Two researchers followed each class (the first author and a research assistant) for the mathematics lessons when the teacher introduced a new topic. Students filled out questionnaires (topic-specific self-efficacy; not reported here, and demographics) prior to the first lesson, as well as short questionnaires at the beginning of each lesson (lesson-specific self-efficacy). In our study, we include four measurement occasions over the course of four lessons (time points 1–4). Each lesson typically lasted 45 min, with the subsequent lesson on the following day (the average time lag between the start of lessons was  $M_{\text{hours}}=30$ , ranging from 45 min to 5 days). The starting point for each sequence of lessons was determined by the introduction to a new topic and ended two to five calendar days later. In Table 1, we provide an overview of the main tasks for each lesson, as described by the teachers, for each of the classes we followed.



**Table 1** Lesson main tasks as described by the teachers

	Lesson 1 main task	Lesson 2 main task	Lesson 3 main task	Lesson 4 main task
Grade 6				
Class 1	To split figures together with your "learning partner"	To draw figures so they become a whole (one)	To compare fractions	To find equal fractions in the fractions game
Class 2	To fold figures in order to understand fractions	To solve tasks regarding fractions from the book	To solve tasks regarding fractions and decimal numbers from the book	To solve tasks regarding fractions and decimal numbers
Class 3	To solve tasks including fractions from the book	To solve tasks on fractions you went through during the lesson	To calculate from a part to a whole	To determine which fraction is greater or smaller
Class 4 <sup>1</sup>	Working in pairs to find different things to weigh, and to use the units kg, hg, and g	Working in pairs to find different things to weigh, and to use the units kg, hg, and g	Working in pairs to solve tasks regarding scale units	Working in pairs to solve tasks regarding scale units
Grade 10				
Class 5	To solve tasks including equations	To solve equations graphically	To solve equations graphically	To solve equations with two unknowns
Class 6	To solve tasks related to proportions	To solve tasks related to proportions	To enter functions in Geogebra	To enter functions in Geogebra
Class 7	To solve equations with two unknowns graphically	To solve equations with two unknowns using the substitution method	To solve equations with two unknowns using the substitution method	(to solve) Tasks regarding finding the weight of screws and nuts
Class 8 <sup>2</sup>	To explore 90 degree triangles in order to understand Pythagoras' theorem	To prove Pythagoras' theorem using Geogebra	To investigate triangles in order to determine whether they were 90 degree angled	No lesson
Class 9	To explain geometrical figures	(to solve tasks related to) Pythagoras' theorem	To use Geogebra in order to solve equations graphically	(to solve tasks related to) Pythagoras' theorem

Notes: 1. Class 4 had two double lessons, where the main topic for lessons 2 and 4 were the same as for lessons 1 and 3, respectively  
 2. Class 8 stayed with the new topic for three lessons only, which means that no data were collected from "Lesson 4" for this class

## 4.2 Participants

Participants were Norwegian students in 6th grade (four classes,  $n=77$ , 41 girls and 36 boys) and 10th grade (five classes,  $n=93$ , 44 girls and 49 boys), approximately 11 and 15 years old, respectively. Our sample was a longitudinal follow-up from a previous study (see Street et al., 2017) from which we followed up 31% of classes. Within the classes that agreed to participate, we obtained an 89% consent rate for student participation. Applying between-group  $t$ -tests, the followed up students from grade 5 to 6 did not differ from those who were not followed up with regard to their recent National Test score in mathematics ( $t_{[186]}=0.48$ ,  $p=0.63$ ) or self-efficacy ( $t_{[184]}=1.73$ ;  $p=0.09$ ). Those who were followed up from grade 9 to 10 outperformed the non-follow-up students ( $t_{[463]}=3.88$ ;  $p<0.001$ ) but did not differ in self-efficacy ( $t_{[454]}=1.26$ ;  $p=0.21$ ). Informed parental/guardian consent was obtained for all students who participated in the study.

Students attended schools with a minimum of 150 students in a sample of municipalities in Norway (see Langfeldt, 2015). The schools varied in terms of prior student performance and socioeconomic background (see Langfeldt, 2015). Norway has a unified rather than stratified school system, where more than 95% of students attend state schools (Statistics Norway, 2021) and ability grouping is illegal (Opplæringslova, 1998, § 8–2). The mixture of ability grouping that characterizes Norwegian classrooms as well as the small degree of systematic differences between Norwegian schools might be advantageous in terms of the generalizability of results.

We collected data from each class for as long as it worked on a new topic in mathematics. Table 2 presents the student and classroom characteristics for our study, including mathematical topic for each class. The total number of data cells was 8,907. Overall missingness was 10.7% (students missing occasions) and 10.6% (missing responses to items). For the current analyses, we included participants who had responded at least twice to the

**Table 2** Student and classroom characteristics

Student characteristics	<i>n</i>	%	M	SD
Gender: Girls	85	50.0		
Boys	85	50.0		
Classroom characteristics	<i>n</i>	%	M	SD
Grade 6	77	45.3		
Age	77		11.71	0.77
Class 1: Fractions	24	14.1		
Class 2: Fractions	23	13.5		
Class 3: Fractions	21	12.4		
Class 4: Measurement	9	5.3		
Grade 10	93	54.7		
Age	93		15.31	0.96
Class 5: Equations	19	11.2		
Class 6: Equations	18	10.6		
Class 7: Equations	24	14.1		
Class 8: Geometry	16	9.4		
Class 9: Geometry	16	9.4		
Lesson reports	<i>n</i>		M	SD
Self-efficacy easy	596		9.12	1.53
Self-efficacy medium	596		8.16	2.06
Self-efficacy hard	596		5.99	2.64

lesson-specific questionnaires. This gave 596 time points (lesson reports) nested in 170 students. The average number of lesson reports was 3.60 (SD=0.60, Range 2–4).

As summarized in Table 2, the topics represented four different content areas: measurement (one grade 6 class), fractions (three grade 6 classes), geometry (two grade 10 classes), and algebra (equations: three grade 10 classes). Given our design, we were able to compare differences in self-efficacy between mathematical content areas within but not between grades. We were, thus, not able to distinguish between effects arising from classroom differences (e.g., teacher effects, student-group composition), grade level (e.g., student age), and potential effects of the content areas. Norwegian schools follow a national curriculum, which outlines the main content areas and learning goals for different grade levels. Given the age of the participants at the point of data collection, the most relevant curriculum to consider is LK06 (valid from 2006–2020; The Norwegian Directorate for Education and Training, 2006). In terms of grade 6, the relevant content areas in the curriculum are numbers and measurement, which are both included from grade 1 as main areas. Looking in more detail, measurement is included in a learning goal at the end of grade 2, while fractions (which falls under numbers) is first introduced as a learning goal at the end of year 4. Considering grade 10, geometry is listed as a main area from 1st through 10th grade, and included in learning goals from the end of grade 2 and up. In contrast, algebra (which equations fall under) is a main area from grade 5 onwards only, and the term equations is first mentioned as a learning goal from the end of grade 7. Overall, the classes focused on topics within content areas which varied in terms of previous exposure in the curriculum.

### 4.3 Procedures and measures

Prior to the introduction of a new topic in mathematics, the first author visited the classes, presented the study, and provided guidance to complete questionnaires. To measure students' lesson-specific self-efficacy in mathematics, we adapted the *Self-Efficacy Gradations of Difficulty* (SEGD; Street et al., 2017). Students completed a brief questionnaire regarding their self-efficacy for learning mathematics *during the lesson* (SEGD lesson) at the beginning of each mathematics lesson in the sequence. Street et al. (2017) confirmed the structural validity of the SEG D (that is, whether the scores of the measure are an adequate reflection of the dimensionality of the construct) and found that students' self-efficacy for easy, medium difficulty, and hard tasks was associated with their scores on a national test in mathematics. The SEG D lesson version measured students' self-efficacy for learning mathematics during the following lesson according to different levels of perceived difficulty, and in relation to five facets for problem solving and self-regulation (see [supplementary materials](#)). Students' expectations were targeted through introductory statements for each of the facets (e.g., *For each of these questions, think about the lesson on <geometry> which is just about to start. How certain are you that you can learn <geometry> with a certain amount of help from the teacher?*). These introductions were followed by statements of varying difficulty (e.g., *if I get a lot of help from the teacher, if I get some help from the teacher, if I get no help from the teacher*). For each item, students were asked to indicate their confidence on a 0–10 scale, with the anchors 0=not at all certain, 5=moderately certain, and 10=completely certain. We investigated the structural validity of the new SEG D lesson measure using Mplus. These analyses confirmed the factor structure of the original SEG D (Street et al., 2017), where a model that took into account students'

**Table 3** Correlation matrix

	1	2	3	4	5			
1. Gender								
2. Grade	0.06							
3. Self-efficacy easy tasks	-0.07	-0.15	***					
4. Self-efficacy medium tasks	-0.02	-0.20	***	0.85	***			
5. Self-efficacy hard tasks	0.04	-0.19	***	0.60	***	0.83	***	
6. Time (lesson)	-0.01	-0.01		0.07		0.11	**	0.16

\*\* Correlation is significant at the 0.01 level (2-tailed)  
 \*\*\* Correlation is significant at the 0.001 level (2-tailed)

self-efficacy for easy, medium difficulty, and hard tasks, with correlated uniquenesses for facets of mathematics, fitted data well (see [supplementary materials](#)).

As shown in Table 3, self-efficacy for easy, medium difficulty, and hard tasks were unrelated with gender, so gender was not included in further analyses. Grade 10 students expressed lower levels of self-efficacy, as seen in the negative associations between Grade and the three self-efficacy constructs. Furthermore, the estimates in Table 3 indicate a trend of positive growth in students’ self-efficacy for medium difficulty and hard tasks, as seen in the positive correlation between Time (lesson) and these constructs.

**4.3.1 Analytic procedures**

Preliminary analyses showed that students’ self-efficacy for easy, medium difficulty, and hard tasks were all stable over time. The mean correlations between adjacent time points for grade 6 and grade 10 were  $r=0.68/0.84$  for easy,  $r=0.76/0.89$  for medium difficulty, and  $r=0.84/0.84$  for hard tasks, respectively. We tested the structural validity of our measures using a series of confirmatory factor analyses (see [supplementary materials](#)). As models fitted data well, we proceeded to setting up our dataset in long format (see Table 4 for two fictive cases) for carrying out multivariate multilevel models for change (MVMLMC). The MVMLMC is an extension of the univariate multilevel model for change, in which

**Table 4** Two fictive cases in long data format (i.e., time points as rows nested within each student)

Student	Class ID	Grade	Lesson	Time	Time <sup>2</sup>	Content area	SE <sub>Easy</sub>	SE <sub>Med</sub>	SE <sub>Hard</sub>
1	1	6	1	0	0	Fractions	9.1	7.8	6.6
1	1	6	2	1	1	Fractions	9.3	8.3	7.0
1	1	6	3	2	4	Fractions	9.4	8.6	7.2
1	1	6	4	3	9	Fractions	9.5	8.9	7.3
2	9	10	1	0	0	Equations	8.6	7.5	5.1
2	9	10	2	1	1	Equations	8.7	7.8	5.8
2	9	10	3	2	4	Equations	8.8	7.9	6.5
2	9	10	4	3	9	Equations	8.9	8.0	6.9

Note for Table 4: Time=the lesson sequence (Lesson 1–4) was recoded into Time (0–3) in order to interpret the intercepts in the MVMLMC as the starting point of the lesson sequence. SE<sub>Easy</sub> self-efficacy for easy tasks, SE<sub>Med</sub> self-efficacy for moderately difficult tasks, SE<sub>Hard</sub> self-efficacy for hard tasks

we investigate individual differences in starting point (i.e., at the beginning, or onset, of a lesson sequence) and change over time in three dependent variables (self-efficacy for easy, medium, and hard tasks) simultaneously.

In the long format, we include unique identifiers for students and classrooms, and covariates (grade and mathematics content area). The coding of the time dimension in the MVMLMC is done at the within-level of the model (i.e., the time points). The "Time" variable starts at zero for interpretation of the starting point (i.e., the latent intercept) in the between part of the model (i.e., students) at the first lesson. Change is specified as a linear trend over time (i.e., a step-wise increase in the dependent variable for each step ahead on the time axis). Time-squared ("Time<sup>2</sup>") was included to pick up nonlinear trends (here quadratic trends), that could capture a steeper acceleration in the beginning and plateauing later. Self-efficacy for tasks of different difficulty levels is included in adjacent columns. This means we can regress each of the three levels of self-efficacy on time and time<sup>2</sup> simultaneously, while allowing the variables to be correlated with each other. Having data in this format allowed us to specify a MVMLMC, in the multilevel structural equation modelling (MSEM) framework (see Fig. 2).

For Fig. 2, we included essential elements only (see [supplementary materials](#) for details). Rectangles depict observed variables. Ellipses depict latent constructs. Single-headed arrows depict directional effects (i.e., regression paths). Double-headed arrows depict either residuals (e.g., for Start SE<sub>Hard</sub> at the within-level), variances (e.g., for Start SE<sub>Hard</sub> at the between-level), or associations. Filled circles ("black dots") depict random effects estimated at the between-level (e.g., the latent intercept Start SE<sub>Hard</sub>, or the latent change Change SE<sub>Hard</sub>). The triangle at the between-level indicates the latent mean (i.e., for latent intercepts and slopes). Covariates are included at the between-level. For each fictive

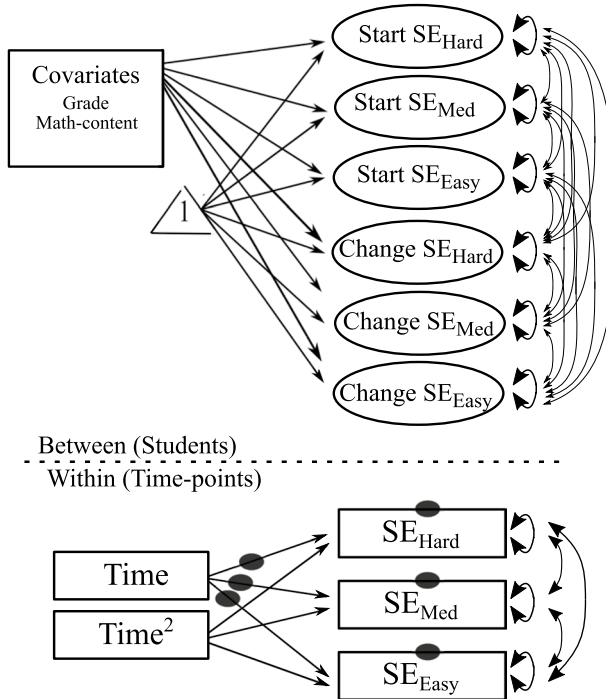


Fig. 2 Multivariate multilevel model for change

student in Table 4, we can imagine three trajectories, a top trajectory depicting self-efficacy for easy mathematics tasks over the four lessons, followed by a middle trajectory (medium difficulty tasks), and a bottom trajectory (hard tasks). Student 1 feels very efficacious for carrying out easy tasks and this sense is sustained and slightly increased over the four lessons (9.1, 9.3, 9.4, and 9.5), while student 2 feels slightly less efficacious than student 1 at the beginning of the first lesson, but also increases this sense slightly across the subsequent lessons. Student 2 has slightly lower self-efficacy for hard tasks compared to student 1, but a steeper increase over time than in the case for student 1. Overall, students are likely to have different starting points of self-efficacy in Lesson 1. This is referred to as individual differences in starting points (intercepts). The students' self-efficacy also changes differently over time. This is referred to as individual differences in change (slopes).

The MVMLMC enables us to investigate many parameters simultaneously. As depicted in Fig. 2, we at the within-level regressed each of the three self-efficacies (for easy, medium, and hard tasks) on Lesson (coded as "Time") and Lesson<sup>2</sup> (coded as "Time<sup>2</sup>"). The linear effect of time (0–1–2–3) captures uniform step-wise increments in self-efficacy over time. The quadratic effect of time (0–1–4–9) captures U-shaped or inverted-U-shaped change in self-efficacy over time. We estimated random intercepts for each self-efficacy variable (indicated with the black dot in Fig. 2). These give three latent intercept variables at the between-level (e.g., Start SE<sub>Hard</sub>). These latent intercepts are interpreted as the starting points of three self-efficacies in the first lesson. The variance of the latent intercept indicates individual differences in starting points for the three levels of self-efficacy (e.g., Start SE<sub>Hard</sub>). We estimate random effects of linear time (indicated with the black dot). These give latent change (or latent slope) constructs at the between-level (e.g., Change SE<sub>Hard</sub>). The variance of this construct indicates individual differences in change for the three levels of self-efficacy (i.e., if students have different self-efficacy trajectories across the four lessons).

### 4.3.2 Modelling strategy

We specified three models. To answer the first and second research questions, we specified a MVMLMC for the three levels of self-efficacy simultaneously. We observed the level of the latent intercepts (for probing Hypothesis 1), latent slopes (Hypothesis 2), and intercept and slope variances (Hypothesis 3). To answer the third research question (Hypothesis 4) we, in the next model, included grade as covariate in the model. To answer the fourth research question (no hypothesis), we included mathematics topic as a dummy-coded covariate in a model for sixth-grade students only (0=fractions, 1=measurement) and tenth-grade only (0=algebra, 1=geometry).

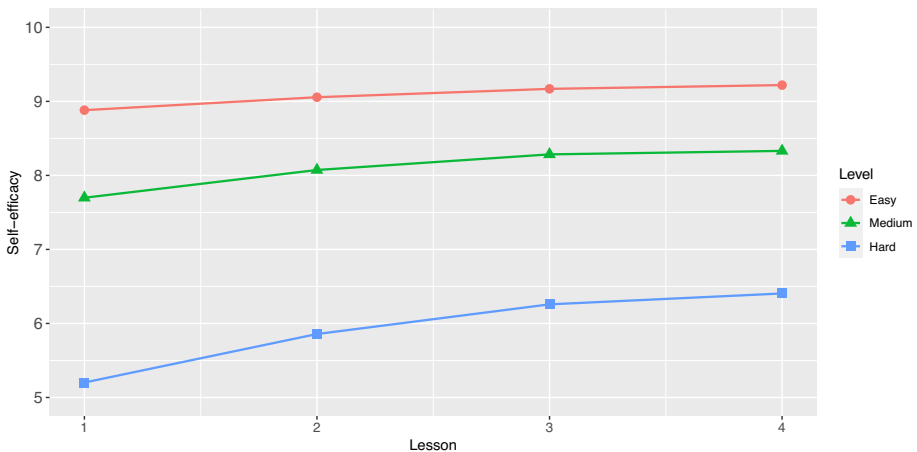
We ran all models using the Bayesian estimator in Mplus 8.6 (Muthén & Asparouhov, 2012; Muthén & Muthén, 2017), using non-informative priors. Bayesian statistics are suitable for models with complex between-level variance parameters such as ours (Muthén & Asparouhov, 2012; Zitzmann et al., 2016). We used the posterior predictive *p*-value (PPP) with values close to 0.5 indicating good model fit for our baselines model, and observing the potential scale reduction (PRS, with values  $\leq 1.05$  indicating appropriate convergence). We report the 95% credibility intervals for parameter estimates from the posterior distribution. Missing data are estimated as default.

## 5 Results

### 5.1 Starting point and change of students' mathematics self-efficacy

In order to answer the first research question “What is the starting point and change of students' mathematics self-efficacy for easy, medium difficulty, and hard tasks during a lesson sequence of a new topic?”, we ran a MVMLMC including students' self-efficacy for easy, medium difficulty, and hard tasks. In terms of self-efficacy at the beginning of the first lesson (starting point), students reported stronger self-efficacy for easier (Start  $SE_{Easy}$ ,  $M=8.88$ , 95% credibility interval [8.63, 9.10]), as compared with medium (Start  $SE_{Med}$ ,  $M=7.70$  [7.39, 8.02]) or harder tasks (Start  $SE_{Hard}$ ,  $M=5.2$ , [4.84, 5.63]; see Fig. 3). The differences were credible when comparing the self-efficacy intercepts of easy with medium tasks ( $\Delta M=1.17$ , [1.02, 1.34]), easy with hard ( $\Delta M=3.67$ , [3.36, 3.99]), and medium with hard ( $\Delta M=2.495$ , [2.285, 2.72]), all mean differences being credible. This provided support for our first hypothesis.

In the case of all three levels of difficulty, there was credible growth (change) across the lessons (Change  $SE_{Easy}$ ,  $B=0.21$ , [0.03, 0.35]; Change  $SE_{Med}$ ,  $B=0.46$ , [0.26, 0.63]; Change  $SE_{Hard}$ ,  $B=0.78$ , [0.50, 1.06]), confirming our second hypothesis. Furthermore, this change was nonlinear in the case of self-efficacy for medium difficulty (quadratic term =  $-0.08$  [ $-0.14, -0.02$ ]) and hard tasks (quadratic term =  $-0.13$ , [ $-0.22, -0.04$ ]). The change in students' self-efficacy (growth) was progressively and credibly steeper for medium than for easy tasks ( $\Delta_{Slope} = .255$ , [.12, .40]), hard than for medium tasks ( $\Delta_{Slope} = .32$ , [.12, .53]), and for hard than for easy tasks ( $\Delta_{Slope} = .58$ , [.33, .88]). Overall, this indicates students' self-efficacy grew credibly over the lesson sequence, where this growth (a) was steeper in relation to harder tasks and (b) plateaued after the first or the second lessons, in relation to medium difficulty and hard tasks. These trends are illustrated visually in Fig. 3.



**Fig. 3** Starting point and change of students' self-efficacy for easy, medium difficulty, and hard tasks



## 5.2 Individual differences in starting point and change of students' mathematics self-efficacy

In order to answer the second research question “Are there individual differences in the starting point and change of students' mathematics self-efficacy for easy, medium difficulty, and hard tasks according to student grade?”, we inspected variances of the intercepts (starting points) and slopes (changes) of students' self-efficacy for easy, medium difficulty, and hard tasks. Our findings were in support of our third hypothesis. In terms of intercepts, there was less individual variance associated with self-efficacy for easy tasks (Start  $SE_{Easy}$ ,  $\sigma^2=2.14$ , [1.635, 2.69]) than for medium difficulty (Start  $SE_{Med}$ ,  $\sigma^2=4.13$ , [3.14, 5.135]) and hard tasks (Start  $SE_{Hard}$ ,  $\sigma^2=6.19$ , [4.76, 7.83]). All three differences in intercept variances were credible: between medium and easy ( $\Delta\sigma^2=1.985$ , [1.23, 2.71]), hard and easy ( $\Delta\sigma^2=4.03$ , [2.76, 5.53]), and hard and medium ( $\Delta\sigma^2=2.07$ , [0.99, 3.16]).

In terms of slope variance, there was broader slope variance for hard (Change  $SE_{Hard}$ ,  $\sigma^2=0.17$ , [0.08, 0.27]), followed by medium difficulty (Change  $SE_{Med}$ ,  $\sigma^2=0.10$ , [0.05, 0.16]), and easy (Change  $SE_{Easy}$ ,  $\sigma^2=0.07$ , [0.035, 0.12]) tasks. The difference was credible in terms of hard versus easy ( $\Delta\sigma^2=.095$ , [.01, .195]) but not for the other differences in slope variances. Overall, we found more individual variation of self-efficacy starting-point and change in association with harder tasks.

## 5.3 Year group differences in starting point and change of students' mathematics self-efficacy

To answer our third research question “What is the starting point and change of students' mathematics self-efficacy for easy, medium difficulty, and hard tasks, according to student grade?”, we included grade as a predictor in our multilevel model for change. Findings from research question 3 are illustrated in Fig. 4. In partial support of our fourth hypothesis, we found grade credibly predicted a lower starting point (intercept) of students' self-efficacy for both medium difficulty ( $B = -0.19$ , [-0.34, -0.05]) and hard tasks ( $B = -0.25$ , [-0.44, -0.05]). This means that students in higher grades had lower self-efficacy starting points for medium difficulty and hard tasks, where these differences were in the order of  $-0.19$  and  $-0.25$  of a point, respectively, on an 11-point scale (see Fig. 4).

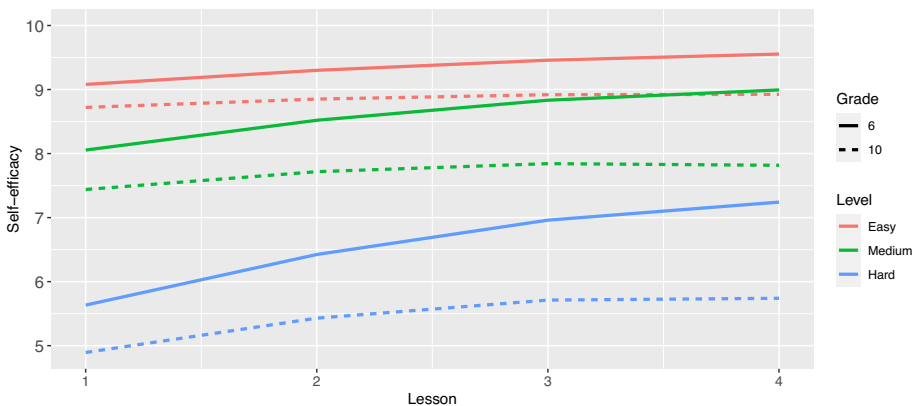


Fig. 4 Year group differences in starting-point and change of students' self-efficacy

Furthermore, we investigated to what degree grade predicted the growth of students' self-efficacy over time. Similar to the effect of grade on self-efficacy starting point, we found student grade negatively predicted the growth (change) of students' self-efficacy for medium difficulty tasks ( $B = -0.62$ ,  $[-1.28, -0.01]$ ). Thus, grade 6 students' self-efficacy grew more steeply than that of grade 10 students across the lesson sequence, in terms of medium difficult tasks (see Fig. 4).

## 5.4 Differences in starting point and change of students' mathematics self-efficacy related to mathematics content areas

In order to answer our final research question "What is the starting point and change of students' mathematics self-efficacy for easy, medium difficulty, and hard tasks, according to mathematics content areas?", we included dummy variables for mathematical content area in two separate models, i.e., differences related to content areas were investigated separately in grade 6 ( $n_{\text{Students}} = 77$ ,  $n_{\text{Time-points}} = 266$ ) and grade 10 ( $n_{\text{Students}} = 93$ ,  $n_{\text{Time-points}} = 330$ ). These models gave us estimates on the effects of focusing on one specific mathematics content area versus another (grade 6: measurement vs fractions; grade 10: geometry vs algebra) on students' self-efficacy starting points and change. For grade 6 students, there was an effect of mathematics content area on the change of students' self-efficacy across the lesson sequence (see Fig. 5). In the three classes that focused on fractions, there was credibly steeper growth in the case of hard tasks ( $B = -0.625$ ,  $[-1.17, -0.04]$ ,  $\beta = -0.41$ ,  $[-0.71, -0.05]$ )), as compared to the class that focused on measurement. There were no differences in change in relation to students' self-efficacy for easy or medium difficulty tasks, nor in self-efficacy starting point for any levels of difficulty. In terms of interpretation, these estimates indicate students' self-efficacy for hard tasks changed by around 2 points more in the fractions classes (on an 11-point scale), as compared with the measurement class, across the sequence of four lessons.

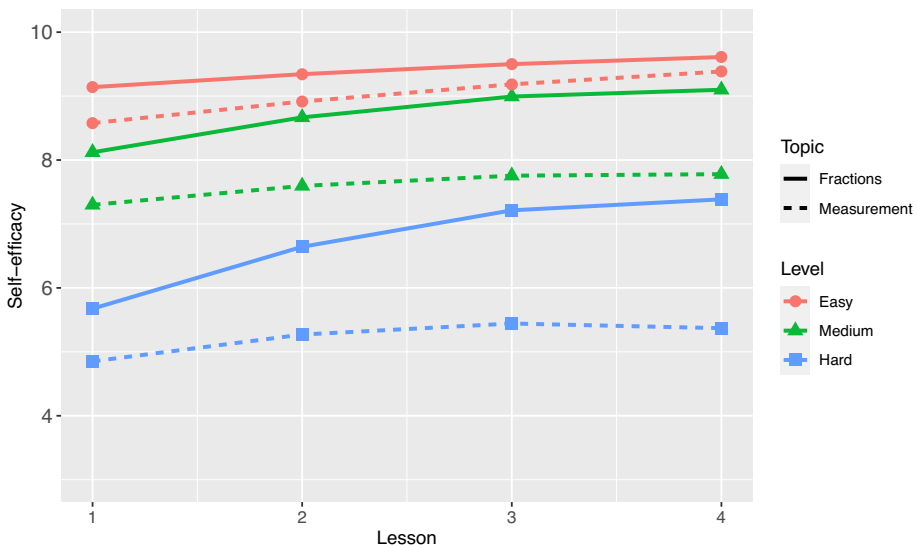
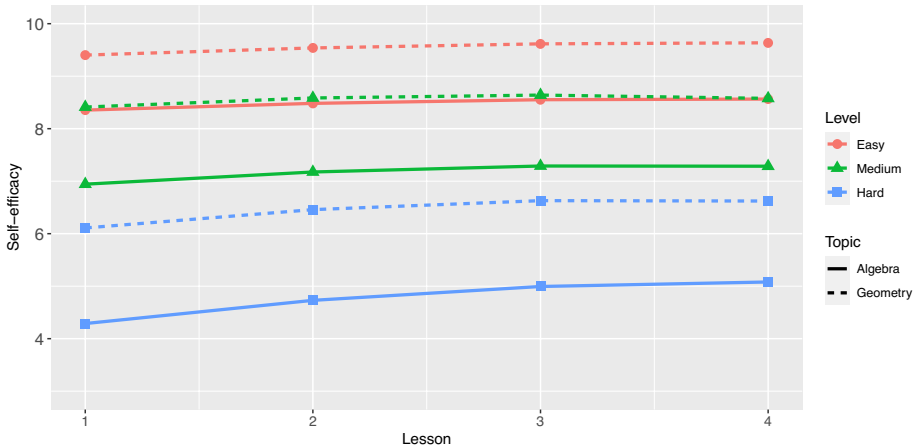


Fig. 5 Starting point and change of self-efficacy according to content area; grade 6



**Fig. 6** Starting point and change of self-efficacy according to content area; grade 10

In terms of grade 10 students, there was an effect of mathematics content area on the starting point of students' self-efficacy at the beginning of the first lesson (see Fig. 6). Across all levels of difficulty, students in the two classes that focused on geometry had stronger self-efficacy at the beginning of lesson 1, as compared with students in the three classes that focused on algebra: self-efficacy for easy tasks ( $B = 1.05$ , [0.31, 1.785],  $\beta = 0.30$  [0.10, 0.49]), medium task ( $B = 1.47$ , [0.45, 2.43],  $\beta = 0.31$  [0.10, 0.49]) and hard tasks ( $B = 1.82$ , [0.66, 2.985],  $\beta = 0.33$  [0.12, 0.51]). This means that the self-efficacy of students in the geometry classes were 1.05, 1.47, and 1.82 points stronger (on an 11-point scale), for easy, medium difficult, and hard tasks, respectively, at the start of the first lesson as compared to students in the algebra classes.

Given the small number of classes included for each mathematics content area, the findings from research question 4 must be interpreted with caution, as we cannot be sure whether differences in starting point and change of self-efficacy are related to differences between the classrooms (e.g., quality of instruction, student characteristics), or to differences in the mathematics content areas. We will discuss in the next section the potential implications from these explorations.

## 6 Discussion

In this paper, we investigated changes in students' difficulty-specific self-efficacy during a sequence of lessons in actual classroom settings, capturing the process of self-efficacy change as grade 6 and grade 10 students were learning a new topic in mathematics. Applying multilevel multivariate models for change, we investigated four research questions, the findings of which we discuss and connect to relevant literature below. After we discuss the findings for each research question in order, we consider limitations of our study.

## 6.1 Stronger self-efficacy for easier tasks, and steeper self-efficacy change for harder tasks

In support of Hypothesis 1 for research question 1, we found higher starting point for easier, as compared with harder, tasks, i.e., students had a stronger sense of mathematics self-efficacy for easier levels of perceived task difficulty. Our findings thus confirm previous studies (e.g., Reinhard & Dickhäuser, 2009; Street et al., 2017) that students consider the difficulty of the task at hand when formulating their self-efficacy. Succeeding on easy tasks requires less skills, effort, and persistence than succeeding on harder tasks; thus, students are likely to report strong self-efficacy in association with easy tasks.

Still in relation to research question 1, we found credible change of self-efficacy across the lessons (Hypothesis 2), with steeper change in association with harder, as compared with easier, tasks. For medium difficulty and hard, but not easy, tasks, there was a non-linear shape to the growth of students' self-efficacy, where initially steeper changes were followed by a flattening trend. To our knowledge, no previous studies investigated change of self-efficacy as a function of perceived task difficulty. It is possible students were cautious in their initial assessment of capability in the case of harder tasks, considering the relative novelty of the mathematical tasks. In that case, self-efficacy might have grown as a function of firsthand experiences with the topic. Also, the lessons might have fostered an increased sense of mastery through developing skills, which may be increasingly important in terms of self-efficacy for tasks of higher difficulty. In the context of a new mathematics topic, mastery experiences during lessons in mathematics appear to be particularly important in the case of harder, as compared with easier tasks, in terms of fostering self-efficacy growth. Our findings are partially consistent with the findings by Borgonovi and Pokropek (2019), who found that exposure on hard tasks was particularly important for the self-efficacy of lower achieving students. Differently to Borgonovi and Pokropek (2019), we used individual students' perceived task difficulty, while Borgonovi and Pokropek aggregated students' confidence ratings to determine easy versus hard tasks. It is likely that perception of task difficulty is relative to achievement level; thus, the "hard tasks" in the study by Borgonovi and Pokropek (2019) might not have been hard enough to uncover an effect of these types of tasks on high achieving students' self-efficacy.

## 6.2 Individual differences in starting point and change of self-efficacy

We found support for Hypothesis 3 in research question 2, predicting credible variance of self-efficacy intercept and slope. Furthermore, the magnitude of these differences increased with level of perceived difficulty. In more detail, there were more individual variations for harder, as compared to easier, tasks in terms of strength of self-efficacy at the beginning of the first lesson. Similarly, in terms of self-efficacy change, there was credibly higher variance in relation to hard tasks, as compared with easy tasks. Individual variations could be related, for example, to student achievement levels, as per the study by Borgonovi and Pokropek (2019). It would be valuable in a future study to unpack the relative contributions to self-efficacy from perceived and objective task difficulty, achievement levels, and mastery experiences (exposure) in mathematics. As discussed in relation to Hypothesis 1, success on harder versus easier tasks requires more in regard to, for example, skills, effort, and perseverance. In terms of implications for classroom practice, our findings suggest the need for adapted support for students,

particularly when working on harder tasks, where differences in terms of mathematics self-efficacy appear larger: both in terms of students' initial self-efficacy when starting a new topic and in terms of the change of their self-efficacy across the subsequent lessons.

### 6.3 Higher starting point and steeper change for younger students' self-efficacy

In relation to research question 3, we found both higher starting-point and steeper change for grade 6, as compared with grade 10, students in terms of medium difficulty tasks, while in terms of starting point, the effect of grade was credible also in relation to students' self-efficacy for hard tasks. Hypothesis 4, regarding group differences in starting point, was thus partially confirmed. Previous studies have reported stronger mathematics self-efficacy for younger students (e.g., Bong, 2009; Zimmerman & Martinez-Pons, 1990a, 1990b), and research on related constructs indicate a developmental decline in mathematics motivation and self-beliefs over time (e.g., Gottfried et al., 2007; Jacobs et al., 2002). Nicholls (1978) discussed how older students are likely more adept at assessing their capabilities due to cognitive maturity as well as performance experiences, which indicates our findings might be a reflection of the older students' weaker, albeit more accurate, self-efficacy.

In terms of students' self-efficacy for medium difficult and hard tasks, the steepest change in self-efficacy was seen at the start of the lesson sequence, followed by a tapering growth. In a previous study the effect of performance on self-efficacy was strongest from the first to the second trial, as compared with subsequent trials (Shea & Howell, 2000). It is possible that students had gained enough experience with the new topic to calibrate their self-efficacy more accurately after one or two lessons (i.e., a process toward a state of self-efficacy equilibrium; see, e.g., Williams & Williams, 2010), and thus additional mastery experiences after this time had little influence on their self-efficacy formation.

In terms of the group differences in growth, it is possible that the older students have developed a more stable sense of self (e.g., Donnellan et al., 2012; Marsh & Grayson, 1994) and that task experiences and feedback are thus less influential in terms of changing their mathematics self-efficacy. While developmental factors such as cognitive maturity and realism might contribute to explain age differences in students' self-efficacy, school and learning contexts might also play a role. Students are subjected to increasing normative comparison in schools (Magro et al., 2020) which may influence the way students perceive their own capabilities. Furthermore, there might be systematic differences between grade 6 and grade 10 classrooms in terms of the kinds of learning experiences available to them. It would be important for future research to include more grades in the comparison to help explore some of the aforementioned hypotheses and paint a more detailed picture of how changes in self-efficacy might play out differently across grades.

### 6.4 Content area differences on starting point and change of students' self-efficacy

When we explored the role of mathematics content area on students' self-efficacy, we found differences in starting point (in relation to grade 10 classes) and change (in relation to grade 6 classes) of students' self-efficacy. In our sample, grade 6 students in the three classes focusing on fractions evidenced steeper change in their self-efficacy for hard tasks, as compared to students in the class focusing on measurement, where the growth in self-efficacy was relatively flat across the four lessons. While across all the classes students were introduced to a new topic, students' familiarity with different aspects of the topic might differ depending on their previous exposure with the particular mathematical

content area as per the mathematics curriculum they go through. Considering the national curriculum context, Norwegian grade 6 students are likely more familiar with the content area of measurement, as compared to fractions. This familiarity could potentially make a difference in the change of their self-efficacy in that it is possible that grade 6 students after repeated exposure had developed a stable sense of their capabilities in measurement, resulting in little change across the lessons. In contrast, students who were introduced to a topic within a relatively less familiar content area might have benefitted more, in terms of self-efficacy growth, from the lessons introducing them to a topic within this content area. This line of reasoning is consistent with the suggestion by Bong (1997) that similarity perceptions are part of the mechanism through which self-efficacy judgments are generalized. Accordingly, we suggest it is possible that familiarity and similarity are related to the stability of self-efficacy over time.

In terms of grade 10, there were no differences in self-efficacy change over time even though there were differences in the starting point self-efficacy: students in the two classes that focused on geometry had stronger self-efficacy beliefs at the start of lesson 1 (for all levels of perceived difficulty), as compared to students in the three classes that focused on algebra. Considering, again, the context of the national curriculum, it is likely grade 10 students were relatively more familiar with aspects of geometry as a content area, as compared with algebra. Besides familiarity, it is also plausible that special characteristics of the content may somehow interfere with students' self-efficacy beliefs. For example, geometry offers a domain where students who may not have been particularly successful with numbers (as in arithmetic and algebra) can feel more confident and use their spatial reasoning and visualization skills to do well. These factors might have made a difference in terms of the strength of students' initial self-efficacy when approaching the topic. The fact that we did not see stronger growth over time in relation to the less familiar content area of algebra (as was the case in grade 6) might be related to the relatively more stable self-efficacy beliefs of the older, as compared with the younger, students (as per RQ3).

It is important to stress that the findings from RQ4 are exploratory, as our study was not initially designed to investigate the effect of content area on the starting point and change of students' self-efficacy. Given the small and uneven number of classrooms used for the comparisons, the findings must be interpreted with appropriate caution. For example, we cannot conclude that the differences are due to the mathematical content area rather than the particular students or the teaching in each of the classrooms, or characteristics of the student group (e.g., average performance level, proportion of disruptive students, etc.). In our study, we showed that the content is a potentially important factor to consider, but the specific way content can affect self-efficacy change over a sequence of lessons we cannot really tell from our study. Our findings highlight the importance of considering the mathematical content area as an important contextual factor when measuring students' self-efficacy in future research, as it is possible that the strength of students' self-efficacy would fluctuate as a result of the content area of the topic of the lessons. We followed classes as the teachers introduced students to "something new," and this ecological approach was not compatible with asking the teachers to focus on a specific content area or topic. It would be important to investigate the role of content area in a more systematic way in a future study by planning, for example, observations of the same classrooms across topics in different content areas, which would help understand better how mathematics content interacts with self-efficacy change.

## 6.5 Limitations

There were a number limitations to our study, some of which we already alluded to earlier. First, as our study was designed for micro-longitudinal starting point and change, we did not focus on between-classroom or between-school differences. Such an investigation would be important to carry out in a study with a substantially larger between-group sample size. Such a study could also incorporate effects of, e.g., individual socioeconomic background and teacher, student group, and school characteristics. Second, in our modelling, we used a discrete time point approach, treating each lesson as a socially defined event in time. Stemming from this, our findings should not be interpreted in relation to the passing of time (linear), but rather the experience of events (i.e., classroom lessons). An important future methodological advancement could be to apply continuous time models (see, e.g., DuBois et al., 2013). Third, the study was carried out in two year groups in Norway. The contents, sequence, and depth of coverage of various topics in the mathematics curriculum, the mixed ability classrooms, and the relative autonomy teachers enjoy for implementing lesson sequences are likely to differ from other countries. It will be valuable to investigate the starting point and change of students' mathematics self-efficacy in larger samples, other grade levels, and within other socio-cultural contexts.

## 7 Concluding remarks

In this study, we investigated both the starting point of and changes in students' self-efficacy across a sequence of lessons in mathematics when students were introduced to a new topic. This is the first study we know of that has investigated self-efficacy change across regular lessons in school mathematics classrooms. Our statistical methodology enabled us to investigate changes in self-efficacy for easy, medium difficult, and hard tasks simultaneously across a sequence of lessons within a mathematical content area.

In our study, we found evidence for mean level changes (particularly early in the learning sequence) as well as individual variations in students' self-efficacy starting point and change. A stable perception of self can be beneficial, in that temporary setbacks will not weaken self-efficacy (Bandura, 1997). At the same time, weak self-efficacy is related to disengagement and anxiety in mathematics (see, e.g., Dowker et al., 2016; Martin et al., 2015), which would be maladaptive over time. Previous studies have highlighted the importance of supporting students' growth mindsets (Dweck, 2002) for learning mathematics (Bonne & Johnston, 2016; Collie et al., 2019; Muis, 2004), as these beliefs are related to students' perceptions of their potential for change. We reiterate these recommendations, and further suggest that the appraisals students make during the initial stages of learning of a new topic are particularly salient. We found evidence of steeper changes in self-efficacy earlier in a learning sequence, which is likely related to students gaining experience with the task or topic, and calibrating (Williams & Williams, 2010) their self-efficacy. Thus, students might particularly benefit from teacher support to appraise their learning experiences (in line with a growth mindset) during these early stages of learning, when their self-efficacy is most likely to change (Usher & Pajares, 2008).

Considering our findings regarding task difficulty, it seems that students' lesson experiences (including, for example, experiences of perseverance, effort, and task success/completion), are more important to the growth of their self-efficacy for harder, relative to



easier, tasks. Previous studies highlighted the importance of memorability to affect changes in students' mathematics self-efficacy (Butz & Usher, 2015; Stylianides & Stylianides, 2014), and several studies emphasize the role of productive struggle for student learning and engagement with mathematics (Bobis et al., 2021; Warshauer, 2015). Higher levels of effort and perseverance might mean that students' work on harder tasks contribute to forming memorable events (Butz & Usher, 2015; Marmur, 2019), to a larger degree than their experiences from working on easy tasks. At the same time, the relationship between task difficulty and self-efficacy is unlikely to be straightforward, as indicated by the interaction effects with achievement levels found per Borgonovi and Pokropek (2019) as well as our own findings regarding credible individual variations in students' starting point and change. Students' beliefs about ability (Dweck, 2002) might interact with the performance–self-efficacy relationship as well as the effects of task difficulty on self-efficacy, through the appraisals students make of their performance experiences. Overall, supporting students' growth mindset in order to instill a willingness to struggle as well as an adaptive appraisal of task experiences, combined with teacher support while working on harder tasks, might be beneficial to support student changes in mathematics self-efficacy.

Our study supports the theoretical tenets that self-efficacy is a dynamic construct and changes in relation to influential events and shifting circumstances. Our findings give empirical evidence of mean-level changes across lessons in mathematics, where the change of students' self-efficacy was characterized by a sharp initial increase, followed by a flattening or stabilizing trend. This was especially typical for self-efficacy for hard tasks, while the growth shape tended to flatten with lower task difficulty. This demonstrates that working on new tasks is particularly salient to students' self-efficacy for hard tasks, and highlights the importance of considering levels of perceived task difficulty when investigating students' self-efficacy. Overall, our micro-analytic approach to mathematics self-efficacy provides an important window into students' learning experiences and raises the issue of whether similar trends would be found in other school subjects such as science (see, e.g., Bong, 1997).

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10649-022-10165-1>.

**Author contribution** All authors contributed to the study conception and design. Material preparation and data collection were performed by Karin E. S. Street. Analysis was performed by Karin E. S. Street and Lars-Erik Malmberg. The first draft of the manuscript was written by Karin E. S. Street and all authors commented on and revised multiple versions of the manuscript. All authors read and approved the final manuscript.

**Funding** Open access funding provided by Western Norway University Of Applied Sciences. This work was supported by the Norwegian Research Council (Norges Forskningsråd); the Grant Number is 218282/H20 (PRAKUT: Learning regions project).

**Data availability** All data and material for the study has been submitted to the Norwegian Centre for Research Data (NSD).

**Code availability** Not applicable.

## Declarations

**Ethics approval** The study was approved by the Norwegian Centre for Research Data (NSD) and the Central University Research Ethics Committee (CUREC) at Oxford University.

**Consent to participate** All participants gave informed, opt-in consent to participate.

**Consent for publication** All participants gave informed, opt-in consent including that the results may be published in an academic article.

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## 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(3), 269–298. <https://doi.org/10.1006/ceps.1996.0926>
- Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman and Co.
- Bernacki, M. L., Nokes-Malach, T. J., & Alevyn, V. (2015). Examining self-efficacy during learning: variability and relations to behavior, performance, and learning. *Metacognition and Learning*, 10(1). <https://doi.org/10.1007/s11409-014-9127-x>
- Bobis, J., Russo, J., Downton, A., Feng, M., Livy, S., McCormick, M., & Sullivan, P. (2021). Instructional moves that increase chances of engaging all students in learning mathematics. *Mathematics*, 9(6), 1–19. <https://doi.org/10.3390/MATH9060582>
- Bong, M. (1997). Generality of academic self-efficacy judgments: evidence of hierarchical relations. *Journal of Educational Psychology*, 89(4), 696–709. <https://doi.org/10.1037/0022-0663.89.4.696>
- Bong, M. (2001). Between- and within-domain relations of academic motivation among middle and high school students: self-efficacy, task-value, and achievement goals. *Journal of Educational Psychology*, 93(1), 23–34.
- Bong, M. (2005). Within-grade changes in Korean girls' motivation and perceptions of the learning environment across domains and achievement levels. *Journal of Educational Psychology*, 97(4), 656–672. <https://doi.org/10.1037/0022-0663.97.4.656>
- Bong, M., & Hocevar, D. (2002). Measuring self-efficacy: multitrait-multimethod comparison of scaling procedures. *Applied Measurement in Education*, 15(2), 143–171. <https://doi.org/10.1207/S15324818AME1502>
- Bong, M. (2009). Age-related differences in achievement goal differentiation. *Journal of Educational Psychology*, 101(4), 879–896. <http://search.proquest.com/docview/614546142?accountid=38978>
- Bonne, L., & Johnston, M. (2016). Students' beliefs about themselves as mathematics learners. *Thinking Skills and Creativity*, 20, 17–28. <https://doi.org/10.1016/j.tsc.2016.02.001>
- Borgonovi, F., & Pokropek, A. (2019). Seeing is believing: task-exposure specificity and the development of mathematics self-efficacy evaluations. *Journal of Educational Psychology*, 111(2), 268–283. <https://doi.org/10.1037/edu0000280>
- Butz, A. R., & Usher, E. L. (2015). Salient sources of early adolescents' self-efficacy in two domains. *Contemporary Educational Psychology*, 42, 49–61. <https://doi.org/10.1016/j.cedpsych.2015.04.001>
- Byars-Winston, A., Diestelmann, J., Savoy, J. N., & Hoyt, W. T. (2017). Unique effects and moderators of effects of sources on self-efficacy: a model-based meta-analysis. *Journal of Counseling Psychology*, 64(6), 645–658. <https://doi.org/10.1037/cou0000219>
- Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: a longitudinal study. *British Journal of Educational Psychology*, 81(1), 78–96. <https://doi.org/10.1348/2044-8279.002004>
- Caprara, G. V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G., Barbaranelli, C., & Bandura, A. (2008). Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning

- in academic continuance and achievement. *Journal of Educational Psychology*, 100(3), 525–534. <http://search.proquest.com/docview/614483692?accountid=38978>
- Chen, P., & Zimmerman, B. J. (2007). A cross-national comparison study on the accuracy of self-efficacy beliefs of middle-school mathematics students. *Journal of Experimental Education*, 75(3), 221–244. <https://doi.org/10.3200/JEXE.75.3.221-244>
- Collie, R. J., Martin, A. J., Bobis, J., Way, J., & Anderson, J. (2019). How students switch on and switch off in mathematics: exploring patterns and predictors of (dis)engagement across middle school and high school. *Educational Psychology*, 39(4), 489–509. <https://doi.org/10.1080/01443410.2018.1537480>
- Dawkins, P. C., & Karunakaran, S. S. (2016). Why research on proof-oriented mathematical behavior should attend to the role of particular mathematical content. *Journal of Mathematical Behavior*, 44, 65–75. <https://doi.org/10.1016/j.jmathb.2016.10.003>
- Donnellan, B. M., Kenny, D. A., Trzesniewski, K. H., Lucas, R. E., & Conger, R. D. (2012). Using trait-state models to evaluate the longitudinal consistency of global self-esteem from adolescence to adulthood. *Journal of Research in Personality*, 46(6), 634–645. <https://doi.org/10.1016/j.jrp.2012.07.005>
- Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics anxiety: what have we learned in 60 years? *Frontiers in Psychology*, 7(508). <https://doi.org/10.3389/fpsyg.2016.00508>
- DuBois, C., Butts, C. T., McFarland, D., & Smyth, P. (2013). Hierarchical models for relational event sequences. *Journal of Mathematical Psychology*, 57(6), 297–309. <https://doi.org/10.1016/j.jmp.2013.04.001>
- Dweck, C. S. (2002). The development of ability conceptions. In Alan Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation*. Academic Press.
- Friedel, J. M., Cortina, K. S., Turner, J. C., & Midgley, C. (2010). Changes in efficacy beliefs in mathematics across the transition to middle school: examining the effects of perceived teacher and parent goal emphases. *Journal of Educational Psychology*, 102(1), 102–114. <http://search.proquest.com/docview/614511223?accountid=38978>
- Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., Oliver, P. H., & Guerin, D. W. (2007). Multivariate latent change modeling of developmental decline in academic intrinsic math motivation and achievement: childhood through adolescence. *International Journal of Behavioral Development*, 31(4), 317–327. <https://doi.org/10.1177/0165025407077752>
- Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/ mathematics performance correspondence. *Journal for Research in Mathematics Education*, 20(3), 261–273.
- Hannula, M. S., Bofah, E. A., Tuohilampi, L., & Metsämuuronen, J. (2014). A longitudinal analysis of the relationship between mathematics-related affect and achievement in Finland. *Proceedings of the Joint Meeting of PME 38 and PME-NA 36, 38*, 249–256.
- Hornstra, L., van der Veen, I., & Peetsma, T. (2016). Domain-specificity of motivation: a longitudinal study in upper primary school. *Learning and Individual Differences*, 51, 167–178. <https://doi.org/10.1016/j.lindif.2016.08.012>
- Jacobs, J. E., Lanza, S., Osgood, D. W., Eccles, J. S., & Wigfield, A. (2002). Changes in children's self-competence and values: gender and domain differences across grades one through twelve. *Child Development*, 73(2), 509–527.
- Johnson, M. L., Edwards, O. V, & Dai, T. (2014). Growth trajectories of task value and self-efficacy across an academic semester. *Universal Journal of Educational Research*, 2(1), 10–18. <https://doi.org/10.13189/ujer.2014.020102>
- Klassen, R. M. (2004). Optimism and realism: a review of self-efficacy from a cross-cultural perspective. *International Journal of Psychology*, 39(3), 205–230. <https://doi.org/10.1080/00207590344000330>
- Klassen, R. M., & Usher, E. L. (2010). Self-efficacy in educational settings: recent research and emerging directions. In T. C. Urdan & S. A. Karabenick (Eds.), *Advances in Motivation and Achievement* (Vol. 16, pp. 1–33). Emerald Group Publishing. [https://doi.org/10.1108/S0749-7423\(2010\)000016A004](https://doi.org/10.1108/S0749-7423(2010)000016A004)
- Krawitz, J., & Schukajlow, S. (2018). Do students value modelling problems, and are they confident they can solve such problems? Value and self-efficacy for modelling, word, and intra-mathematical problems. *ZDM - Mathematics Education*, 50(1–2), 143–157. <https://doi.org/10.1007/s11858-017-0893-1>
- Küchemann, D., & Hoyles, C. (2001). Investigating factors that influence students' mathematical reasoning. *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education*, 265–272.
- Langfeldt, G. (2015). Skolens kvalitet skapes lokalt. In *Skolens kvalitet skapes lokalt. Presentasjon av funn fra forskningsprosjektet "Lærende Regioner."* Fagbokforlaget.
- Liu, Q., Liu, J., Cai, J., & Zhang, Z. (2020). The relationship between domain- and task-specific self-efficacy and mathematical problem posing: a large-scale study of eighth-grade students in China. *Educational Studies in Mathematics*, 105(3), 407–431. <https://doi.org/10.1007/s10649-020-09977-w>

- Magro, S. W., Fraley, R. C., & Roisman, G. I. (2020). Early maternal sensitivity and teacher–student relationship quality across grade school: enduring or transient associations? *Child Development, 91*(4), e883–e901. <https://doi.org/10.1111/cdev.13322>
- Marmur, O. (2019). Key memorable events: a lens on affect, learning, and teaching in the mathematics classroom. *Journal of Mathematical Behavior, 54*. <https://doi.org/10.1016/j.jmathb.2018.09.002>
- Marsh, H. W., & Grayson, D. (1994). Longitudinal stability of latent means and individual differences: A unified approach. *Structural Equation Modeling: A Multidisciplinary Journal, 1*(4), 317–359. <https://doi.org/10.1080/10705519409539984>
- Marshall, M. A., & Brown, J. D. (2004). Expectations and realizations: the role of expectancies in achievement settings. *Motivation and Emotion, 28*(4), 347–361. <https://doi.org/10.1007/s11031-004-2388-y>
- Martin, A. J., Way, J., Bobis, J., & Anderson, J. (2015). Exploring the ups and downs of mathematics engagement in the middle years of school. *Journal of Early Adolescence, 35*(2), 199–244. <https://doi.org/10.1177/0272431614529365>
- Muis, K. R. (2004). Personal epistemology and mathematics: a critical review and synthesis of research. *Review of Educational Research, 74*(3), 317–377.
- Muthén, B. & Muthén, L. (1998–2017). *Mplus User's Guide* (8th ed.) Los Angeles, CA: Muthén and Muthén.
- Muthén, B. O., & Asparouhov, T. (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychological Methods, 17*, 313–335. <https://doi.org/10.1037/a0026802>
- 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*(3), 800–814.
- Opplæringslova*, (1998) (testimony of Ministry of Education and Research). <https://lovdata.no/lov/1998-07-17-61/§8-4>
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543–578.
- Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24*(2), 124–139. <https://doi.org/10.1006/ceps.1998.0991>
- Phan, H. P. (2012). The development of English and mathematics self-efficacy: a latent growth curve analysis. *The Journal of Educational Research, 105*(3), 196–209. <https://doi.org/10.1080/00220671.2011.552132>
- Reinhard, M.-A., & Dickhäuser, O. (2009). Reinhard & Dickhäuser, 2009.pdf. *Journal of Personality and Social Psychology, 96*(5), 1062–1076.
- Shea, C. M., & Howell, J. M. (2000). Efficacy-performance spirals : an empirical test. *Journal of Management, 26*(4), 791–812.
- Statistics Norway. (2021). *Pupils in primary & lower secondary schools, December 2020*. 14.07.2021
- Street, K. E. S., Malmberg, L.-E., & Stylianides, G. J. (2017). Level, strength, and facet-specific self-efficacy in mathematics test performance. *ZDM - Mathematics Education, 49*(3), 379–395. <https://doi.org/10.1007/s11858-017-0833-0>
- Stylianides, A. J., & Stylianides, G. J. (2014). Impacting positively on students' mathematical problem solving beliefs: an instructional intervention of short duration. *The Journal of Mathematical Behavior, 33*, 8–29. <https://doi.org/10.1016/j.jmathb.2013.08.005>
- Talsma, K., Schütz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa): a meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences, 61*(January), 136–150. <https://doi.org/10.1016/j.lindif.2017.11.015>
- The Norwegian Directorate for Education and Training. (2006). *National Curriculum in Mathematics (MAT1–01)*.
- Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751–796.
- Warshauer, H. K. (2015). Productive struggle in middle school mathematics classrooms. *Journal of Mathematics Teacher Education, 18*(4), 375–400. <https://doi.org/10.1007/s10857-014-9286-3>
- Wigfield, A., Eccles, J. S., & Rodríguez, D. (1998). The development of children' s motivation in school contexts. *Review of Research in Education, 23*, 73–118.
- Williams, T., & Williams, K. (2010). Self-efficacy and performance in mathematics: reciprocal determinism in 33 nations. *Journal of Educational Psychology, 102*(2), 453–466. <https://doi.org/10.1037/a0017271>
- Zimmerman, B. J., & Martinez-Pons, M. (1990a). Student differences in self-regulated learning : relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*(1 March 1990a), 51–59. <https://doi.org/10.1037/0022-0663.82.1.51>
- Zimmerman, B. J., & Martinez-Pons, M. (1990b). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology, 82*(1), 51–59. <https://doi.org/10.1037/0022-0663.82.1.51>

---

Zitzmann, S., Lüdtke, O., Robitzsch, A., & Marsh, H. W. (2016). A Bayesian approach for estimating multilevel latent contextual models. *Structural Equation Modeling*, 23(5), 661–679. <https://doi.org/10.1080/10705511.2016.1207179>

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