TamPub This document has been downloaded from TamPub.uta.fi
The Institutional Repository of University of Tampere

Learning and Instruction 56 (2018) 73–83

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Consistency, longitudinal stability, and predictions of elementary school students' task interest, success expectancy, and performance in mathematics

Katariina Nuutila[a,*], Heta Tuominen[a], Anna Tapola[a], Mari-Pauliina Vainikainen[a,b], Markku Niemivirta[a,c,**]

[a] Faculty of Educational Sciences, University of Helsinki, P.O. Box 9, 00014, Finland
[b] Faculty of Education, University of Tampere, 33014, Finland
[c] Faculty of Educational Sciences, University of Oslo, P.O. Box 1092, Blindern 0317, Oslo, Norway

ABSTRACT

This study examined cross-task consistency and longitudinal stability in elementary school students' task interest, success expectancy, and performance from fourth to sixth grade, and their predictive effects on sixth-grade intrinsic value, self-concept, and achievement in mathematics. The results demonstrated consistency in interest, success expectancy, and performance across tasks and stability over time, and these to predict domain-specific motivation and achievement. Virtually no evidence for reciprocal effects was found for task-specific measures, as only previous task performance predicted change in later success expectancy. Cross-lagged effects were observed, however, for predictions of task motivation and performance on domain-specific motivation and achievement, so that success expectancy predicted intrinsic value, interest predicted self-concept, and task performance predicted both self-concept and achievement. Based on the findings, it would seem that students' task-related motivational experiences are associated with their domain-specific beliefs, and that those, in turn, are to some extent manifested in students' task motivation.

1. Introduction

Students' incentives for engaging in learning activities and the way they perceive their competence are important motivational precursors of achievement outcomes, including school performance (e.g., grades; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005) and educational choices (e.g., choosing non-compulsory courses; Simpkins, Davis-Kean, & Eccles, 2006). These effects also seem to apply to performance in specific tasks (e.g., task interest and self-efficacy in a problem-solving task; Niemivirta & Tapola, 2007) and to achievement in different subject areas (e.g., value and self-concept in reading; Schoor, 2016), among younger students (Eccles & Wigfield, 1993) as well as older students (Guo, Parker, Marsh, & Morin, 2015). Longitudinal studies have been conducted on the development of domain-specific motivation (i.e., students' relatively stable motivational beliefs in relation to a subject domain such as mathematics; Jacobs, Lanza, Osgood, Eccles, & Wigfield, 2002) and its relations to achievement (Arens et al., 2017; Seaton, Marsh, Parker, Craven, & Yeung, 2015), but regarding task-specific motivation, research seems to have focused on single tasks or situations (Ainley, 2006). Relatively few studies have looked at the consistency of task motivation across tasks or stability over time, or its predictions on domain-specific motivation and achievement within a longer time span (Fryer, Ainley, & Thompson, 2016; Rotgans & Schmidt, 2011). The available work has mainly focused on out-of-school settings (e.g., extra-curricular courses) and older students (e.g., Knogler, Harackiewicz, Gegenfurtner, & Lewalter, 2015). As it is often argued that domain-specific motivation (e.g., intrinsic value and self-concept in mathematics) accumulates through repeated experiences in tasks and situations that reflect certain subject areas and related activities (Bong & Skaalvik, 2003; Hidi & Renninger, 2006), it would seem reasonable to investigate whether students' task-specific motivation generalizes across different tasks, and whether they predict similar experiences and domain-specific motivation over time.

Accordingly, the aim of the present study was to examine i) the consistency of students' interest, success expectancy, and performance across different tasks (i.e., cross-task consistency) and over time (i.e., longitudinal stability), ii) their longitudinal reciprocal relationships, and, iii) their predictions on intrinsic value, self-concept, and

* Corresponding author. Faculty of Educational Sciences, University of Helsinki, P.O. Box 9, 00014, Finland.
** Corresponding author. Faculty of Educational Sciences, University of Oslo, P.O. Box 1092, Blindern 0317, Oslo, Norway.
E-mail addresses: katariina.nuutila@helsinki.fi (K. Nuutila), heta.tuominen@helsinki.fi (H. Tuominen), anna.tapola@helsinki.fi (A. Tapola), mari-pauliina.vainikainen@uta.fi (M.-P. Vainikainen), markku.niemivirta@helsinki.fi, markku.niemivirta@iped.uio.no (M. Niemivirta).

https://doi.org/10.1016/j.learninstruc.2018.04.003
Received 3 April 2017; Received in revised form 5 April 2018; Accepted 9 April 2018
Available online 01 May 2018
0959-4752/ © 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/BY/4.0/).

achievement in mathematics. The effects of gender were also taken into account.

1.1. The dynamics of task motivation and performance

Task motivation refers to students' motivational states (e.g., interest, boredom; e.g., Ainley, 2006) and self-appraisals (Zimmerman, 2000) in relation to a specific task such as calculations in a mathematics class. In this study, we will focus on students' task interest and task-specific success expectancy,[1] which are known to have both independent (Lee, 2009; Linnenbrink-Garcia et al., 2010) and joint effects (Niemivirta & Tapola, 2007) on task performance. Task interest (e.g., Graham, Tisher, Ainley, & Kennedy, 2008; also referred to as situational interest; Ainley & Hidi, 2002) is best described as a momentary state of heightened attention and enjoyment triggered by the interaction between the characteristics of the task and the student, which could be short-lived or last until finishing the task. Success expectancy, in turn, is defined as a person's belief about their success in a task (Eccles et al., 1983; related to, and often equivalently used as self-efficacy, which is defined as an individual's confidence in being able to orchestrate and execute actions required for achieving intended results; Bandura, 1986).

With regard to task interest and its outcomes, it is often intuitively assumed that higher interest should enhance performance, but the empirical findings have been mixed. While task interest appears in some cases to enhance task performance (Vainikainen, Salmi, & Thuneberg, 2015), memorizing and text comprehension (Hidi, 2001), and persistence (Thoman, Sansone, & Pasupathi, 2007), interest evoked by seductive details may also hinder performance (Wang & Adesope, 2016). Then again, it is also possible that task interest does not predict task performance directly, but through students' involvement or engagement in the task (Rotgans & Schmidt, 2011). Task performance does not seem to predict subsequent interest independently either (Shen, Chen, & Guan, 2007), although they often are correlated (Jansen, Lüdtke, & Schroeders, 2016). Also, while prior knowledge can facilitate the triggering of task interest (Logtenberg, van Boxtel, & van Hout-Wolters, 2011; Niemivirta & Tapola, 2007), it can also diminish it, if the task is perceived as too easy (Rotgans & Schmidt, 2014).

As to the relationship between success expectancy and task performance, the limited findings available suggest it to be reciprocal, although the empirical evidence is surprisingly sparse (see, Williams & Williams, 2010). This is nevertheless in line with the theoretical assumptions (Schunk & DiBenedetto, 2016), and also echoes the findings from the reciprocal effects model as applied to domain-specific motivation (Marsh & Martin, 2011). The underlying logic here would be that performances viewed as successful raise success expectancy, and perceived failures, by contrast, lower success expectancy. Success expectancy, in turn, enhances performance, as individuals with high success expectancy tend to apply more sustained effort, and they appear to be able to do so even in the face of boredom or other distractions (e.g., Honicke & Broadbent, 2016).

In empirical research, interest and success expectancy have often been found to be correlated, and one explanation for this is that some level of certainty or confidence in one's competence is a prerequisite for interest to arise in the first place (e.g., Bandura, 1986; Silvia, 2003), and positive self-appraisals are likely to enhance positive emotions and to diminish negative feelings (Chen et al., 2016; Tanaka & Murayama, 2014). Then again, positive emotions and interest may boost effort and commitment (Patall, Vasquez, Steingut, Trimble, & Pituch, 2016), thus resulting in experiences of success, which, in turn, may amplify sense of efficacy (Bandura, 1978). The causal relationships between interest, success expectancy, and performance are not clear, although findings from studies looking at their longitudinal relationships seem to support reciprocal effects, with competence perceptions having a stronger effect on achievements and interest than vice versa (Marsh et al., 2005). Other findings, however, complicate the picture. For example, in a study on the dynamics of task-specific motivation, Niemivirta and Tapola (2007) found that in addition to cross-sectional relations, also the changes across the task in interest and self-efficacy were strongly correlated (i.e., change in interest paralleled the change in self-efficacy), and, even more importantly, that the initial level of self-efficacy (but not the change in it) and the change in interest (but not the initial level of it) independently predicted task performance. In a recent study by Fryer et al. (2016) on university students' second language learning, self-efficacy positively predicted competence after the course (i.e., performance in a standardized language test), as did course interest, but task interest at the end of the course did not. Initial self-efficacy also predicted change in task interest over an eight-month period. Linnenbrink-Garcia, Patall, and Messersmith (2013), instead, found in their study on a summer science course that triggered situational interest (i.e., temporary affective state evoked by the context; Hidi & Renninger, 2006) predicted self-efficacy at the end of the course, whereas maintained situational interest (i.e., sustained involvement in and enjoyment of the task content) did not.

In a more recent study, Chen et al. (2016) investigated the predictive effects between situational interest (triggered and maintained) and self-efficacy during a ten-day science intervention. The results showed these to be reciprocally related, and maintained interest to have stronger effect on post-intervention self-efficacy than triggered interest. However, neither of the above two studies measured actual performance or achievement.

1.2. The cross-task consistency and longitudinal stability in task motivation and performance

Although students' domain-specific motivation seems rather stable over time (Spinath & Steinmayr, 2008), it is less clear whether and to what extent the same applies to task motivation. It is often assumed, at least implicitly, that psychological experiences during a task, such as interest, are specific to the task and thereby ‘unique’ and transient (Renninger & Hidi, 2011). Simultaneously, however, it would also be reasonable to argue that students' domain-specific motivational beliefs and other individual tendencies provide the motivational lenses through which one perceives the different situations (Boekaerts & Niemivirta, 2000), thus increasing the likelihood of consistency and stability in task motivation, particularly within a specific subject domain.

Current findings on task interest suggest there to be both task-specificity and cross-task consistency. Students' interest (i.e., mean level) does fluctuate during a task (Niemivirta & Tapola, 2007) as well as across different tasks (Graham et al., 2008) and learning activities (Palmer, 2009; Rotgans & Schmidt, 2011). However, there also seems to be relatively high stability in interindividual differences (i.e., rank-order between students) within and across tasks, at least when the time span between the measurements is relatively short (e.g., tasks are done within one learning session; Rotgans & Schmidt, 2011). As most existing studies have focused on a rather short time span, more is to be learned about the stability in task-related interest over extended periods of time.

Knogler et al. (2015) directly addressed the question of cross-situational consistency and situation-specific variability in interest across different learning activities (e.g., information gathering and role play) over a time period of three weeks. Compared to previous findings (e.g., Rotgans & Schmidt, 2011), they detected less consistency in situational interest, and that even this observed stability was largely explained by

[1] Studies examining the expectancy and value components of motivation often focus on conceptually different but related constructs. Thereby, even though our empirical focus is on individuals' task-specific experiences of interest (e.g., “*The task was interesting*”; Durik & Harackiewicz, 2007) and expectations of success (e.g., “*I performed well in the task*”; Seegers, van Putten, & de Brabander, 2002), we will in our review also consider studies investigating other similar constructs, and thus generalize from these results in order to draw conclusions about previous findings.

domain-specific interest (i.e., individual interest in energy supply). However, as the authors note, the differences in results might partly be due to the longer time period and rather different types of learning activities used. Also, the activities were not representative of any school subjects or content domains as such, but, instead, were built around one broad theme (i.e., energy supply). In the study by Fryer et al. (2016), the tasks used were more similar over time, which might explain their finding of higher stability in task interest. Then again, students' initial task interest was also partly dependent on their prior domain-specific interest.

Research on cross-task consistency or longitudinal stability in success expectancy is still sparse, which may be due to success expectancy being conceived as highly task-specific, and thus sensitive to variation (Zimmerman, 2000). Theoretically, success expectancy should exhibit increasing consistency and stability over time, as students develop more stable and elaborated domain-specific beliefs of their skills and competences, which are then reflected in task situations within a domain (Schunk & DiBenedetto, 2016). However, the empirical evidence for this is surprisingly limited, with only few studies demonstrating weak to moderate stability in task- (Ackerman, Kanfer, & Goff, 1995; Phan, 2014) and course-specific competence perceptions (Galyon, Blondin, Yaw, Nalls, & Williams, 2012; Lane, Hall, & Lane, 2004) over time. Moreover, these studies have focused mainly on older students in out-of-school contexts.

1.3. The relations between task- and domain-specific motivation and achievement

While task motivation is connected to a specific task and the outcomes of the task, domain-specific motivation represents a more stable and less situation-bound set of motivational beliefs and tendencies related to a certain subject area, such as mathematics or reading (Wigfield et al., 1997). Following the expectancy-value theory (Wigfield & Eccles, 2000), intrinsic value (i.e., perceived importance of and the interest towards a domain) and self-concept (i.e., individuals' perception about their own abilities and competences) represent those “want” and “can” aspects of domain-specific motivation that have been consistently found to promote achievement and other educational outcomes (e.g., Guo et al., 2015).

The four-phase model of interest development (Hidi & Renninger, 2006) provides a view on how task motivation might contribute to domain-specific motivation through cumulative experiences, by describing how task interest could lead to more established individual interest (i.e., a tendency to re-engage with the object of interest across time and contexts; Renninger, 2009) via different developmental phases. In its initial phase, interest is momentary and mainly characterised by enjoyment and heightened attention, but as the individual further engages in the given domain, interest develops through different phases into more stable individual interest, associated with personal meaning, knowledge, and value. Thus, although there is a distinction between value beliefs and individual interest as such (Renninger & Hidi, 2016), it would seem reasonable to assume that the cumulative experiences of task-related interest that leads to more stable individual interest might result in stronger intrinsic value as well (Schiefele, 2009).

Similar to the development of individual interest, it has also been proposed that academic success expectancies provide one cognitive basis for developing academic self-concept. The argument is that through repeated exposure to achievement situations with similar tasks, students “develop an aggregated sense of their own academic capability on the basis of salient success or failure experiences” (Bong & Skaalvik, 2003, p. 31). However, to date, no empirical studies have explicitly tested this proposition within a developmental design.

Correlational studies have shown both task interest and individual interest (Durik & Matarazzo, 2009), and self-efficacy and self-concept (Ferla, Valcke, & Cai, 2009) to be associated with each other, but studies looking at their longitudinal predictions are few. In the Fryer et al. (2016) study, task interest was found to predict course interest, which, in turn, predicted domain-specific interest, whereas the Linnenbrink-Garcia et al. (2013) study showed both triggered and maintained interest to predict later individual interest. Chen et al. (2016), in contrast, observed that maintained but not triggered situational interest predicted later individual interest. These findings thus partly support the assumption of task interest reinforcing further individual interest. As to similar studies concerning task-related success expectancies and domain-specific self-concept, we are not aware of any studies that would have investigated these predictions in a longitudinal setting.

Besides the developmental continuums between task interest and intrinsic value, and success expectancy and self-concept, the extent to which these constructs are mutually related is somewhat unclear. Some short-term longitudinal research shows predictions from situational interest to domain-specific self-efficacy (Linnenbrink-Garcia et al., 2013) and from self-efficacy to individual interest (Chen et al., 2016), but these have not been found in longer-term studies (e.g., Fryer et al., 2016). Regarding domain-specific motivation, cross-sectional correlational research demonstrates reciprocality between intrinsic value and self-concept (Liou, 2017), but findings from longitudinal studies are mixed. Some studies show stronger effects from self-concept to interest than vice versa (Marsh et al., 2005), some from interest to self-concept (Ganley & Lubienski, 2016), while others show no effects at all (Spinath & Steinmayr, 2008). As the differences in study designs as such do not seem to explain these inconsistencies, we cannot make any strong conclusions about these effects.

While task motivation is likely to enhance task performance, and also to contribute to domain-specific motivation, research on its long-term influence on academic achievement is limited. Evidence shows success expectancy to be a strong predictor of not only concurrent but also later achievement (Kriegbaum, Jansen, & Spinath, 2015), and these effects to become stronger with age (Chen, Yeh, Hwang, & Lin, 2013). Regarding task interest, only the predictions on immediate task performance have usually been examined. In these studies, the effects have often been relatively small or even non-existent (Ainley, Hillman, & Hidi, 2002; Graham et al., 2008; Zhu et al., 2009). Task interest has, however, been found to be predictive of students' achievement-related behaviour (e.g., engagement; Linnenbrink-Garcia et al., 2013) and indirectly also achievement (Rotgans & Schmidt, 2011), but these mediating effects have been examined within a short time span (e.g., a learning session). Finally, course interest has been found to be predictive of course choices within longer periods of time (e.g., during undergraduate studies over several years) but not achievement as such (Harackiewicz, Barron, Tauer, & Elliot, 2002).

2. The present study

In this study, we examined students' task motivation and performance in terms of their consistency across tasks, stability over time, and relations to domain-specific motivation and achievement through the following research questions: 1) To what extent is there consistency in elementary school students' interest, success expectancy, and performance a) across different types of mathematics tasks, and b) over time? 2a) How are task-related interest, success expectancy, and performance longitudinally related, and 2b) how do they predict intrinsic value, self-concept, and achievement in mathematics?

Studies on the consistency of task motivation are still few, and the available ones have mainly focused on interest (Knogler et al., 2015; Rotgans & Schmidt, 2011). Unlike these studies, which mostly investigate consistency and situation-specificity across quite heterogeneous types of learning activities not representative of any specific subject areas in out-of-school contexts or in the form of an intervention, our study focused on students' task motivation in relation to a specific school subject with tasks that are different in content (e.g., arithmetics

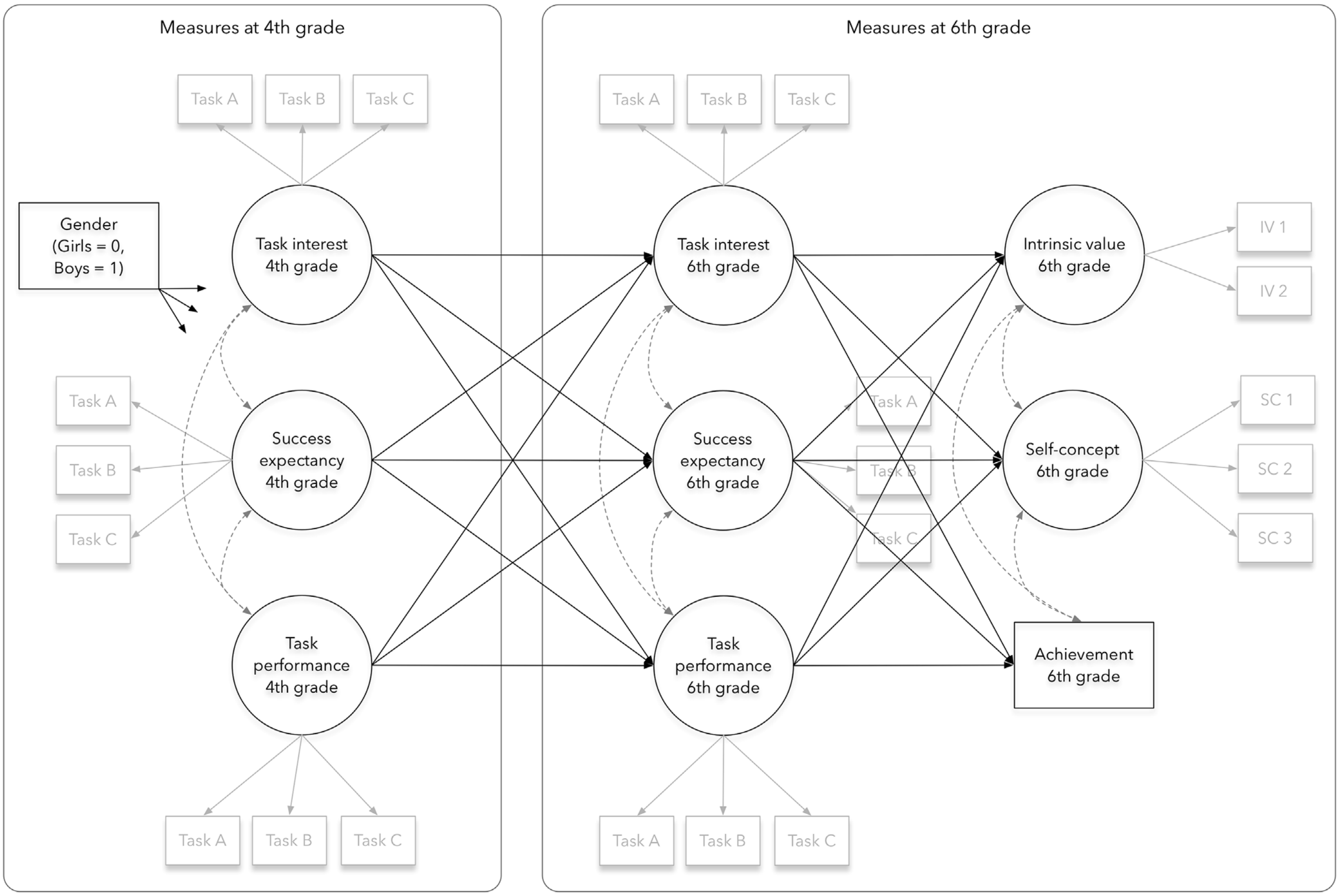

Fig. 1. The hypothetical model of cross-lagged longitudinal predictions. For clarity, arrows representing gender effects, autocorrelations, and residual correlations are omitted.

vs. mental calculation), but represent the same domain (i.e., mathematics), and in which the type of the activity itself is similar across the tasks (i.e., mathematics tasks in the form of a paper and pencil test). Further, our participants, being younger and not from a selective group of students (cf., Knogler et al., 2015; Rotgans & Schmidt, 2011), were more representative of a heterogeneous classroom. The time span of our design (i.e., two years, from fourth to sixth grade) was also relatively long.

The participating students completed a set of three mathematics tasks (tasks A, B, and C) twice, first in the fourth and again in the sixth grade. The tasks differed in content in that they represented different mathematical subdomains and required different mathematical processes (e.g., mental calculation and reasoning). For each task in each occasion, they also rated their task motivation in terms of interest and success expectancy. To address consistency across tasks, we specified a model where both types of task-related motivation and performance across the three tasks were each explained by separate latent factors (see Fig. 1). The rationale was that if the model with latent factors explaining variation in students' task interest, success expectancy, and performance for each set of tasks within each measurement point fit the data, this would indicate congruence across ratings and performance beyond the mere situations. Longitudinal stability and reciprocal effects would then be examined by testing the predictions across the same latent constructs two years later.

In addition to cross-task consistency and longitudinal stability, we also examined the predictions of task-related motivation on domain-specific motivation (i.e., intrinsic value and self-concept) and achievement in mathematics. Such predictions might shed some light on the role task motivation possibly plays in the formation of domain-specific motivation. Studying these relationships during the late elementary school years would seem of special importance, since this time period is crucial for the development and differentiation of domain-specific motivation (Wigfield et al., 1997). The hypothetical model illustrating the specified latent structures and the relationships between all constructs within and over time is depicted in Fig. 1.

Based on previous theorizing (Bong & Skaalvik, 2003) and empirical findings (Arens et al., 2017; Kriegbaum et al., 2015; Seaton et al., 2015), we expected task-related success expectancy and performance to predict each other and likely also self-concept and achievement, and task interest to predict intrinsic value in mathematics. Although grounding on rather weak and inconsistent evidence, we nevertheless further anticipated task interest to predict success expectancy and self-concept (Chen et al., 2016; Ganley & Lubienski, 2016; Linnenbrink-Garcia et al., 2013), and success expectancy to predict task interest and intrinsic value (Chen et al., 2016; Fryer et al., 2016).

Finally, as previous research suggests that particularly mathematics is a subject area where differences between boys and girls might be significant in terms of both interest and perceived competence (Fredricks & Eccles, 2002), we wanted to control for the effects of gender. Despite girls' generally more positive academic motivation and higher achievement, in mathematics they tend to display lower intrinsic value and less positive self-concept in comparison to boys (Guo et al., 2015). However, these effects might also be somewhat specific to the country and educational systems. For example, while in many countries boys outperform girls in mathematics, this does not seem to be the case in Finland (OECD, 2016). Again, the late elementary school years seem to be of particular importance, since the gender gap in math-related self-perceptions seems to increase especially during this time period

(Huang, 2013). Within the present context, then, we expected to find gender differences favouring boys in task- and domain-specific motivation (Spinath, Eckert, & Steinmayr, 2014), but not in performance.

3. Method

3.1. Participants and procedure

The data came from a longitudinal study with students from 24 different elementary schools in Finland (for a description of the project the data were drawn from, see Vainikainen, 2014). The participants included in this study were a total of 865 fourth-graders (52% girls, $M_{age} = 9.60$, $SD = 0.52$) from 51 different classrooms, who completed three sets of mathematics tasks in autumn 2010. After finishing each set, the students rated their task interest and success expectancy. The design was repeated in the sixth grade in spring 2013 ($M_{age} = 12.27$, $SD = 0.46$), and in this occasion, the students also reported their domain-specific motivation (i.e., intrinsic value and self-concept) in mathematics. The final sixth-grade mathematics grade obtained from the class teacher was used as an indicator of students' mathematics achievement.

Before the data collection, students' guardians were informed about the study, and the students were assured of anonymity. Both assessments were conducted by class teachers according to written instructions.[2] The students filled out the test booklets during ordinary classes. The mathematics tasks were embedded into the overall data collection of the project including a variety of tasks and a questionnaire, with four separate 45-minute sessions allocated for the assessment in the fourth grade, and one 90-minute session for the sixth-grade assessment.

3.2. Measures

3.2.1. Task performance and achievement

Three tasks were used to measure performance in different content areas in mathematics. The respective tasks were otherwise identical at the fourth and sixth grades, but due to age-graded differences, they were calibrated so that some easy items were replaced with more difficult items in the sixth-grade versions.

In Task A, Invented Mathematical Concepts, an arithmetical operator is conditionally defined depending on the value of the digits they combine (e.g., *if a > b, lag stands for subtraction, and else for multiplication*). The task was a modified group-version of Sternberg's Triarchic Test (H-version) Creative Number Scale (Sternberg, Castejon, Prieto, Hautamäki, & Grigorenko, 2001), comprising eight items with four multiple-choice alternatives, coded dichotomously for the whole equation. Alphas for the task were 0.68 (fourth grade, 8 items) and 0.76 (sixth grade, 8 items).

In Task B, Hidden Arithmetical Operators (Demetriou, Platsidou, Efklides, Metallidou, & Shayer, 1991), there were one to four hidden operators in each item (e.g., [(*5 a 3*) *b 4 = 6. In this task letter a/b stands for: addition* (+)/subtraction (−)/*multiplication* (×)/*division* (÷)?]). The items were coded dichotomously for a correct answer for all 1–4 operators in the item. Alphas for the task were 0.60 (fourth grade, 8 items) and 0.67 (sixth grade, 7 items).

Task C, Mental Arithmetics, was based on the Arithmetic subscale of the Wechsler Adult Intelligence Scale - Revised (WAIS-R; Wechsler, 1981). The teacher read aloud a mathematical problem (e.g., *If you buy two bus tickets and one ticket costs 3 euros 50 cents, how much money do you get back if you give 10 euros?*) after which the students wrote their answer in their test booklets. The items were coded dichotomously as correct or incorrect. Alphas for the task were 0.66 (fourth grade, 5 items) and 0.75 (sixth grade, 8 items).

The final mathematics grade in the sixth grade was used as a measure of mathematics achievement. In the Finnish school system, grades range from 4 (fail) to 10 (distinction).

3.2.2. Task interest and success expectancy

After each task (A, B, and C), the students rated their task interest and success expectancy similarly in the fourth and sixth grades. Single-item measures with a 7-point Likert-scale (1 = Not at all true - 7 = Very true) were used. The items were: *"This task was very interesting"* (Task Interest), and *"I believe I did very well on this task"* (Success Expectancy).[3]

3.2.3. Intrinsic value and self-concept in mathematics

In the sixth grade, in addition to the measures of mathematics performance and task motivation, the students reported their intrinsic value and self-concept in mathematics. Both intrinsic value (e.g., *"I am interested in mathematics"*, 2 items) and self-concept (e.g., *"I am good at mathematics"*, 3 items) were measured with a 7-point Likert-scale (1 = Not at all true - 7 = Very true). The composite reliability was 0.73 for intrinsic value and 0.91 for self-concept.

3.3. Overview of analyses

The data were analyzed using Mplus Statistical Software (Muthén & Muthén, 1998–2017). A cross-lagged longitudinal panel design within the structural equation modeling (SEM) framework was used to test the measurement models and predictions as specified according to our hypothetical model (see Fig. 1). As some minor indications of clustering across classes were detected by means of intraclass correlations (e.g., design effects being slightly above 2; see Muthén & Satorra, 1995), we took this into account by using a more appropriate estimation (i.e., through the TYPE = COMPLEX specification as implemented in the Mplus statistical software). In this approach, the standard errors using a sandwich estimator and Chi-square test of model fit are calculated in a manner that takes into account the non-independence of observations due to clustering of the sample.

In the model, task-related interest, success expectancy, and performance were first specified as latent variables assumed to cause variation in their indicators (interest, success expectancy, and performance in the three different task sets, respectively). As to longitudinal predictions, sixth-grade task interest, success expectancy, and performance were regressed on fourth-grade task interest, success expectancy, and performance. Similarly, intrinsic value, self-concept, and achievement in mathematics were regressed on both fourth- and sixth-grade measures of task-related interest, success expectancy, and performance. Corresponding observed measures over time as well as the within-task residual covariances were let to correlate. Gender (coded as girls = 0, boys = 1) was specified to predict all latent factors and final mathematics achievement.

To evaluate the model fit, we applied three absolute model fit indices, the Chi-square test of model-fit, Root Mean Square Error of Approximation (RMSEA; cutoff-value close to < 0.06; Steiger, 1990), and Standardized Root Mean Square Residual (SRMR; cutoff-value close to < 0.08; Hu & Bentler, 1999), and two comparative fit indices, the Comparative Fit Index (CFI; cutoff-value close to > 0.95; Bentler, 1990) and Parsimonious Comparative Fit Index (PCFI; cutoff-value close to > 0.50; Mulaik et al., 1989). All solutions were generated using maximum likelihood estimation with robust standard errors (MLR), and missing data were handled with full-information maximum likelihood method (e.g., Dong & Peng, 2013).

[2] In the Finnish elementary schools (grades 1–6; ages 7–12), instruction is usually given by the same class teacher in most subjects (Finnish National Agency for Education, 2017).

[3] Due to the aims and nature of the underlying longitudinal project our data were drawn from, post-task single item measures were used in order to introduce minimal interference with the actual tasks.

Table 1
Descriptive statistics for all variables included in the model.

Item/variable	M	SD	Range	Skewness	Kurtosis
Grade four (t1)					
Task A					
Task performance	5.16	2.00	0–8	–.31	–.60
Task interest	4.44	2.25	1–7	–.34	−1.33
Success expectancy	5.05	1.97	1–7	–.86	–.42
Task B					
Task performance	5.05	1.38	0–8	−1.12	1.59
Task interest	4.00	2.09	1–7	.00	−1.25
Success expectancy	4.20	2.00	1–7	–.23	–.97
Task C					
Task performance	1.72	1.49	0–5	.52	–.80
Task interest	3.25	2.01	1–7	.44	–.99
Success expectancy	4.02	1.89	1–7	–.14	–.98
Grade six (t2)					
Task A					
Task performance	4.68	2.22	0–8	–.01	−1.06
Task interest	3.76	2.07	1–7	.056	−1.29
Success expectancy	4.20	1.87	1–7	–.30	–.92
Task B					
Task performance	2.35	1.53	0–7	.47	–.24
Task interest	2.77	1.88	1–7	.76	–.62
Success expectancy	2.98	1.80	1–7	.56	–.70
Task C					
Task performance	4.52	2.24	0–8	–.21	–.86
Task interest	2.95	1.70	1–7	.56	–.66
Success expectancy	4.01	1.70	1–7	–.09	–.78
Mathematics					
Intrinsic value	5.31	1.36	1–7	–.88	.45
Self-concept	4.77	1.50	1–7	–.54	–.33
Achievement	8.19	1.11	4–10	–.36	–.37

Note. Task-specific variables represent single items, and intrinsic value and self-concept represent composite scores.

4. Results

We first estimated a model as shown in Fig. 1, with good fit to the data, χ^2 (206, $N = 865$) = 495.201, $p < .001$; CFI = 0.96; PCFI = 0.66; RMSEA = 0.04; SRMR = 0.04. We then trimmed the model by removing non-significant predictive paths, with no particular improvement in the model fit, χ^2 (223, $N = 865$) = 508.526, $p < .001$; CFI = 0.96; PCFI = 0.71; RMSEA = 0.04; SRMR = 0.04. The only difference between the original and trimmed model was the change of the prediction from sixth-grade task performance on sixth-grade self-concept from non-significant to significant. For illustration purposes, descriptive statistics for all variables are given in Table 1, and correlations in Appendix. Factor loadings, latent correlations, and predictive effects from the SEM model are presented in Fig. 2.

Our first objective (1a) was to examine the consistency of motivation and performance across three different mathematics tasks. The good model fit allows us to conclude that the model with single latent factors representing variation in students' respective ratings of interest and success expectancy as well as their performance across different tasks at both measurement points was adequate. Standardized factor loadings on each factor were all significant and relatively similar in magnitude, ranging from 0.54 to 0.72 ($p < .001$) for task interest, from 0.52 to 0.74 ($p < .001$) for success expectancy, and from 0.52 to 0.72 ($p < .001$) for performance.

As to our second objective (1b), the examination of longitudinal stability in students' task motivation and performance, we found all successive predictions to be significant. Highest stability was found for task performance ($\beta = 0.83$, $t = 16.06$, $p < .001$), followed by task interest ($\beta = 0.43$, $t = 8.19$, $p < .001$) and success expectancy ($\beta = 0.31$, $t = 5.90$, $p < .001$).

Our next objectives focused on (2a) mutual within-measurement correlations and across-measurement predictions between task-related

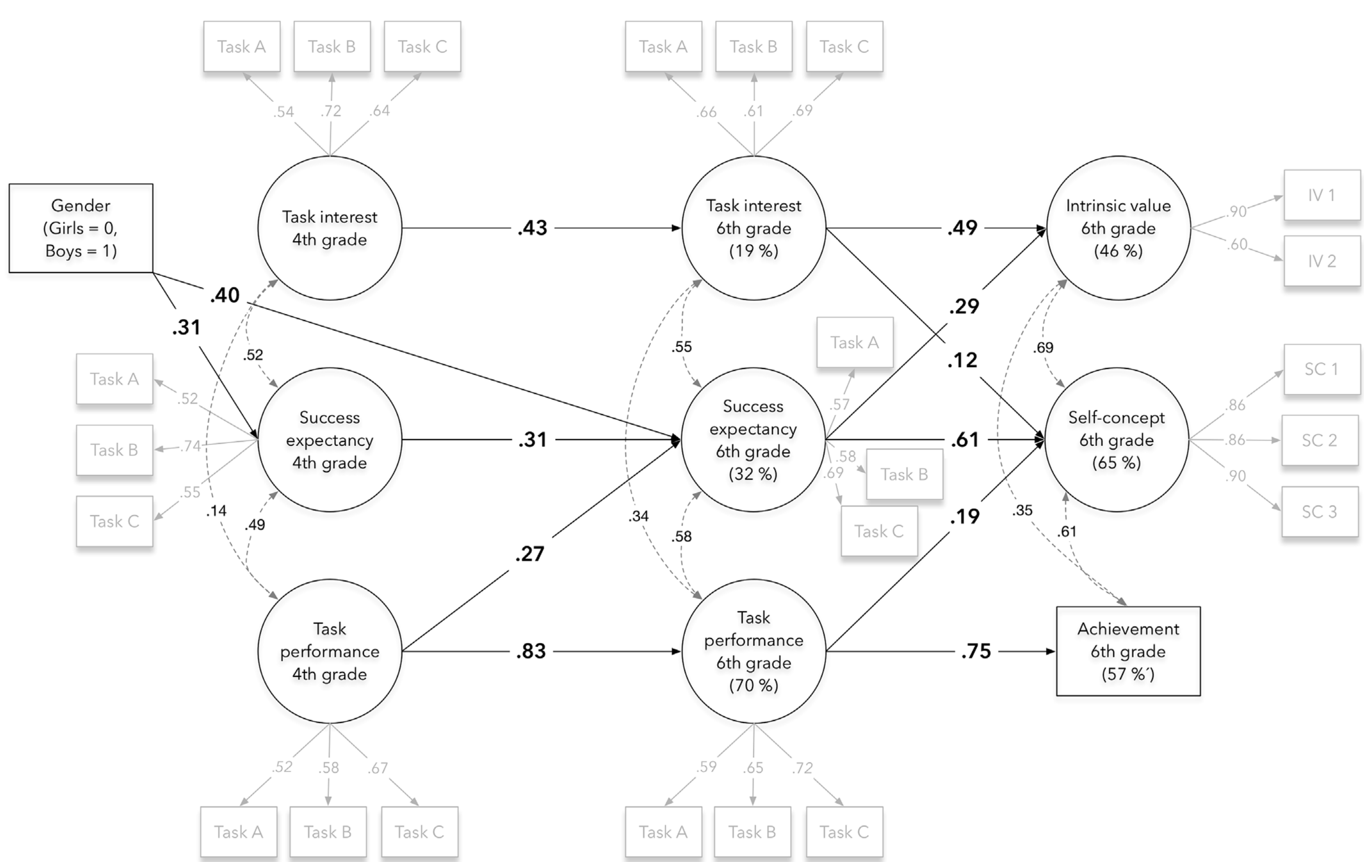

Fig. 2. The empirical model of cross-lagged longitudinal predictions. Only significant (standardized) direct effects and correlations are displayed.

Table 2
Standardized total and total indirect effects of task interest, success expectancy, task performance, and gender on intrinsic value, self-concept, and achievement in mathematics.

	Intrinsic value t2			Self-concept t2			Achievement t2		
	β (se.)	*t*	*p*	β (se.)	*t*	*p*	β (se.)	*t*	*p*
Total effects									
Task interest t1	.21 (.03)	6.44	< .001	.05 (.02)	2.82	.005			
Success expectancy t1	.09 (.02)	4.10	< .001	.19 (.04)	5.07	< .001			
Task performance t1	.08 (.03)	2.96	.003	.32 (.05)	6.22	< .001	.63 (.05)	12.99	< .001
Gender	.14 (.07)	1.96	.050	.39 (.07)	5.58	< .001	.10 (.09)	1.10	.276
Total indirect effects									
Gender	.20 (.07)	2.87	.004	.36 (.07)	5.36	< .001	.11 (.07)	1.55	.122

Note. Gender coded as: girls = 0, boys = 1.

interest, success expectancy, and performance, as well as (2b) their effects on domain-specific motivation and achievement. As shown in Fig. 2, task interest, success expectancy, and performance all significantly correlated with each other at both occasions. Strong associations were observed between task interest and success expectancy, $r = 0.52$ and $r = 0.55$ at fourth and sixth grades, respectively. Similarly high were correlations between success expectancy and performance at both measurement points, $r = 0.49$ and $r = 0.58$, respectively. The correlation between task interest and performance was significant but modest in the fourth grade ($r = 0.14$), but stronger in the sixth grade ($r = 0.34$). Regarding mutual predictions between task interest, success expectancy, and performance, only the effect of fourth-grade task performance on the change in success expectancy in the sixth grade was significant ($\beta = 0.27$, $t = 4.32$, $p < .001$).

Similar to the task-specific relations, the associations between intrinsic value, self-concept, and achievement were all significant. As assumed, the correlations between intrinsic value and self-concept ($r = 0.69$), and self-concept and achievement ($r = 0.61$) were high, while the correlation between intrinsic value and achievement ($r = 0.35$) was smaller but moderately high. In line with our expectations, effects from task motivation and performance on domain-specific motivation and achievement showed sixth-grade task interest to be a strong predictor of intrinsic value ($\beta = 0.49$, $t = 11.03$, $p < .001$). It also had a small yet significant effect on self-concept ($\beta = 0.12$, $t = 3.26$, $p < .001$). Sixth-grade success expectancy was a strong predictor of self-concept ($\beta = 0.61$, $t = 11.62$, $p < .001$), and had a moderate effect on intrinsic value ($\beta = 0.29$, $t = 4.88$, $p < .001$). Sixth-grade task performance had a significant but rather small effect on self-concept ($\beta = 0.19$, $t = 4.12$, $p < .001$), and a strong effect on mathematics achievement ($\beta = 0.75$, $t = 21.69$, $p < .001$).

Regarding indirect effects, we found fourth-grade task interest to predict intrinsic value ($\beta = 0.21$, $t = 6.44$, $p < .001$) and self-concept ($\beta = 0.05$, $t = 2.82$, $p = .005$) via sixth-grade task interest, and fourth-grade success expectancy to predict self-concept in mathematics ($\beta = 0.19$, $t = 5.07$, $p < .001$) and intrinsic value ($\beta = 0.09$, $t = 4.10$, $p < .001$) via sixth-grade success expectancy. Fourth-grade task performance had a significant indirect effect on intrinsic value ($\beta = 0.08$, $t = 2.96$, $p = .003$) via sixth-grade success expectancy, and on self-concept via sixth-grade success expectancy ($\beta = 0.16$, $t = 4.27$, $p < .001$) and task performance ($\beta = 0.16$, $t = 4.03$, $p < .001$), and on achievement in mathematics ($\beta = 0.63$, $t = 13.00$, $p < .001$) via sixth-grade task performance.

As to gender effects, we found boys to report higher success expectancy than girls at fourth grade ($\beta = 0.31$, $t = 2.90$, $p = .004$) and more positive (or less negative) change in it at grade six ($\beta = 0.40$, $t = 4.92$, $p < .001$), after controlling for the effect of prior task performance. In addition to the direct effects, gender indirectly predicted domain-specific motivation through success expectancy so that boys reported higher intrinsic value ($\beta = 0.12$, $t = 3.72$, $p < .001$) and displayed more positive self-concept ($\beta = 0.25$, $t = 4.74$, $p < .001$) in mathematics. All total and total indirect effects are reported in Table 2.

In total, our model explained 19% ($p < .001$) of the variance in task interest, 32% ($p < .001$) in success expectancy, 70% ($p < .001$) in task performance, 46% ($p < .001$) in intrinsic value, 65% ($p < .001$) in self-concept, and 57% ($p < .001$) in achievement.

5. Discussion

The aim of this study was to investigate cross-task consistency and longitudinal stability in task interest, success expectancy, and performance, as well as their predictions on intrinsic value, self-concept, and achievement in mathematics.

The results demonstrated congruence in students' interest and success expectancy ratings and performances across the three different types of mathematics tasks at both measurement points (i.e., at grades four and six), thus indicating relatively high cross-task consistency. This consistency, in the form of common latent factors explaining the variation in students' interest and success expectancy across the task, could be taken to reflect something that is shared across the tasks despite their independence, such as interest and confidence in doing mathematics tasks in general. Thus, although not directly observed, it would seem safe to suggest that students' task-related motivational appraisals were not just a bottom-up function of the task, but also influenced by the domain-related motivational beliefs they bring to the situation (cf., Boekaerts & Niemivirta, 2000).

In a sense, the longitudinal stability we observed in students' task motivation is in agreement with this. Considering that we measured interest and success expectancy in relation to matching tasks over a time span of two years, the observed consistency is likely to represent such stability in students' experiences that goes beyond the mere situation. This, in turn, is supported by the relatively strong predictions of task motivation on intrinsic value and self-concept in mathematics. In other words, the common components of interest and success expectancy across the different tasks significantly contributed to individual differences in parallel domain-specific motivation. Although the proximity of the measures here is likely to partly explain the strong effects observed, it would seem reasonable to argue that task-specific appraisals of interest and success expectancy are strongly linked with more general self-beliefs within the same domain. This inference is further qualified by the indirect effects of fourth-grade task interest on intrinsic value and fourth-grade success expectancy on self-concept. In fact, the majority of the variance in intrinsic value and self-concept was explained by the preceding task interests and success expectancies. Although not directly addressed, these findings are in agreement with the cumulative developmental processes explicated in the four-phase model of interest (Hidi & Renninger, 2006) and alike theorizing on self-efficacy and self-concept (Bong & Skaalvik, 2003). While our findings do not substantiate these theories, they do not contradict them either, thus adding another empirical piece to the puzzle.

In addition to consistency across tasks and stability over time, we were also interested in the mutual predictions between task motivation and performance. Our assumptions of reciprocal effects were not supported, since the only cross-lagged effect found was from fourth-grade performance on sixth-grade success expectancy, thus indicating that higher task performance in grade four contributed to an increase in success expectancy in grade six. This goes against the reciprocal effects model consistently supported in studies on domain-specific motivation (Arens et al., 2017; Seaton et al., 2015), but concurs with findings suggesting that among younger students, achievement is likely to steer competence perceptions rather than vice versa (Skaalvik & Hagtvet, 1990).

Interest and success expectancy predicted neither each other nor later task performance, despite their significant within-measurement correlations at both waves. This also is somewhat unanticipated, considering theoretical presumptions (Jacobs et al., 2002) as well as previous evidence on task-specific motivation (Chen et al., 2016; Fryer et al., 2016; Kriegbaum et al., 2015; Linnenbrink-Garcia et al., 2013; Niemivirta & Tapola, 2007; Tanaka & Murayama, 2014) and corresponding findings on domain-specific motivation (Ganley & Lubienski, 2016; Liou, 2017; Marsh et al., 2005; Seaton et al., 2015). The limited variance to be explained due to high stability might justify part of it, but solely in relation to task performance, since the stability of interest and success expectancy over time was only moderate.

The cross-lagged predictions of sixth-grade task motivation on intrinsic value and self-concept were, in contrast, more in line with our expectations. Success expectancy predicted higher intrinsic value, and task interest, in turn, predicted more positive self-concept, as was anticipated based on previous findings (Ganley & Lubienski, 2016; Liou, 2017; Marsh et al., 2005). Task performance predicted both achievement and self-concept, as found in previous studies (Viljaranta, Tolvanen, Aunola, & Nurmi, 2014). That is, of students' domain-specific evaluations about the intrinsic value of mathematics and their competence in it, only the latter was associated with task performance, yet the actual competence was relatively strongly predicted by it. Again, it is possible that the above relations are partly due to the proximity of the measurements, but it could nevertheless also suggest that students' motivational beliefs in relation to a subject domain are manifested in task motivation, but are less dependent on individual task performances. Our finding of the indirect effect of fourth-grade task performance on mathematics achievement, and on self-concept through increased success expectancy at grade six, is in line with this.

Last, as girls' math-related motivation is often lower than boys', even in the absence of competence discrepancy (Guo et al., 2015), gender effects were also taken into account. Regarding motivation, our results mostly corroborated previous findings (Spinath et al., 2014): no differences were found in task interest and performance, yet boys reported higher success expectancy, which, in turn, mediated similar effects on intrinsic value and self-concept. Also, the change in success expectancy over time was more positive (or less negative) among boys, which reinforces the notion that the gender gap in competence perceptions may increase during the late elementary school years (Huang, 2013).

When interpreting our results, some limitations need to be considered. First, it should be noted that we used single-item post-task measures for assessing task motivation due to the need to minimize interference with the actual task performance. As both single-item and post-task measures have been successfully used in previous studies (Ainley & Patrick, 2006; Durik & Harackiewicz, 2007; Gogol et al., 2014; Silvia, 2003), this should not be taken as a serious threat to the validity of the findings. However, pre-task measures of task motivation would undoubtedly be more optimal, as such measures might more appropriately address the influence of motivational appraisals on subsequent task performance and be less subject to bias. Due to calibrating the difficulty of the tasks to different ages, the tasks were not entirely identical, which prevented the analysis of true developmental change in performance over time. This clearly would have been desirable. Also, parallel measures of domain-specific motivation at the fourth grade would have been informative.

To conclude, our findings demonstrate consistency in students' task interest and success expectancy across different tasks within the same subject domain, and stability over an extended period of time. However, no support was found for the reciprocal effects model, with limited evidence favouring the notion of achievement contributing to motivation rather than vice versa (cf. the skill development model; Calsyn & Kenny, 1977). Expected cross-lagged predictions were only observed between the more proximate measures of task interest and success expectancy and intrinsic value and self-concept. The results provide some, although mostly indirect, support for the notion that students' task motivation is not just a function of the task, but also partly reflects their domain-specific motivational beliefs, and that task-specific motivational experiences contribute to students' domain-specific beliefs. As a practical implication of this, we see our findings suggesting that apart from nourishing students' existing domain-specific motivational beliefs, supporting the triggering of interest and experiencing confidence in relation to everyday learning tasks in ordinary classroom situations might also hold importance.

In order to better capture the developmental dynamics, future research on task motivation and performance as well as on domain-specific motivational beliefs and achievement should include more frequent measurements with shorter time span in between. For an even more thorough view on the fluctuations of task motivation and its influence on performance, these could further include repeated measures within each task (cf. Niemivirta & Tapola, 2007). Finally, research should also be expanded to other school subjects and age groups, as the emerging patterns of predictions and change might be different from those found in the present study.

Acknowledgements

This research was supported by grants from the Academy of Finland (#287170 to the second author and #279742 to the last author), and the Jenny and Antti Wihuri Foundation (#160423 to the third author). The first author has a doctoral student position at the University of Helsinki Doctoral Programme in Psychology, Learning and Communication. We would like to thank the reviewers for constructive comments, members of the Motivation, Learning, and Well-being research group for support, and all students and teachers for participation.

Appendix. Correlations for all variables

		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.	18.	19.	20.	21.
1.	A Task interest t1	–																				
2.	B Task interest t1	.33**	–																			
3.	C Task interest t1	.39**	.46**	–																		
4.	A Task interest t2	.21**	.17**	.18**	–																	
5.	B Task interest t2	.18**	.21**	.16**	.45**	–																
6.	C Task interest t2	.22**	.26**	.28**	.48**	.41**	–															
7.	A Success expectancy t1	.41**	.11**	.17**	.13**	.03	.09*	–														
8.	B Success expectancy t1	.15**	.37**	.18**	.12**	.13**	.11**	.37**	–													
9.	C Success expectancy t1	.13**	.24**	.40**	.12**	.13**	.14**	.33**	.36**	–												
10.	A Success expectancy t2	.11**	.11**	.11**	.59**	.25**	.29**	.21**	.17**	.15**	–											
11.	B Success expectancy t2	.12**	.10**	.06	.22**	.52**	.25**	.13**	.15**	.17**	.39**	–										
12.	C Success expectancy t2	.07	.11**	.11**	.23**	.22**	.36**	.17**	.27**	.25**	.37**	.39**	–									
13.	A Task performance t1	.18**	.07	.03	.12**	.00	–.01	.38**	.23**	.03	.21**	.12**	.15**	–								
14.	B Task performance t1	.05	.03	–.02	.09*	.03	.04	.21**	.35**	.05	.15**	.14**	.19**	.38**	–							
15.	C Task performance t1	.00	.04	.02	.07	.10	.02	.07	.24**	.31**	.10	.23**	.33**	.27**	.36**	–						
16.	A Task performance t2	.12**	.07	.03	.38**	.14**	.19**	.20**	.17**	.07	.49**	.18**	.29**	.33**	.29**	.32**	–					
17.	B Task performance t2	.17**	.10*	.08	.23**	.26**	.19**	.20**	.25**	.13**	.29**	.38**	.30**	.28**	.30**	.26**	.48**	–				
18.	C Task performance t2	.02	.03	–.04	.12*	.09	.13**	.18**	.25**	.15**	.23**	.21**	.55**	.35**	.33**	.51**	.43**	.46**	–			
19.	Intrinsic value t2	.13**	.18**	.15**	.39**	.33**	.38**	.07*	.14**	.11**	.30**	.23**	.31**	.21**	.12**	.21**	.20**	.24**	.29**	–		
20.	Self-concept t2	.09*	.14**	.13**	.35**	.30**	.34**	.13**	.32**	.21**	.45**	.41**	.53**	.24**	.22**	.39**	.34**	.39**	.48**	.59**	–	
21.	Achievement t2	.04	.03	–.03	.21**	.14**	.14**	.11**	.28**	.13**	.22**	.21**	.36**	.30**	.38**	.47**	.42**	.48**	.59**	.33**	.58**	–

Note. A = Task A; B = Task B; C = Task C; t1 = Time 1; t2 = Time 2.
** p < .01, * p < .05.

References

Ackerman, P. L., Kanfer, R., & Goff, M. (1995). Cognitive and noncognitive determinants and consequences of complex skill acquisition. *Journal of Experimental Psychology: Applied, 1*, 270–304. http://dx.doi.org/10.1037/1076-898X.1.4.270.

Ainley, M. (2006). Connecting with learning: Motivation, affect and cognition in interest processes. *Educational Psychology Review, 18*, 391–405. http://dx.doi.org/10.1007/s10648-006-9033-0.

Ainley, M., & Hidi, S. (2002). Dynamic measures for studying interest and learning. In P. R. Pintrich, & M. L. Maehr (Vol. Eds.), *Advances in motivation and achievement: New directions in measures and methods: Vol. 12*, (pp. 43–76). Amsterdam: JAI.

Ainley, M., Hillman, K., & Hidi, S. (2002). Gender and interest processes in response to literary texts: Situational and individual interest. *Learning and Instruction, 12*, 411–428. http://dx.doi.org/10.1016/S0959-4752(01)00008-1.

Ainley, M., & Patrick, L. (2006). Measuring self-regulated learning processes through tracking patterns of student interaction with achievement activities. *Educational Psychology Review, 18*, 267–286. http://dx.doi.org/10.1007/s10648-006-9018-z.

Arens, A. K., Marsh, H. W., Pekrun, R., Lichtenfeld, S., Murayama, K., & vom Hofe, R. (2017). Math self-concept, grades, and achievement test scores: Long-term reciprocal effects across five waves and three achievement tracks. *Journal of Educational Psychology, 109*, 621–634. http://dx.doi.org/10.1037/edu0000163.

Bandura, A. (1978). The self system in reciprocal determinism. *American Psychologist, 33*, 344–358.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice-Hall.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238–246. http://dx.doi.org/10.1037/0033-2909.107.2.238.

Boekaerts, M., & Niemivirta, M. (2000). Self-regulation in learning: Finding a balance between learning and ego-protective goals. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.). *Handbook of self-regulation* (pp. 417–450). San Diego, CA: Academic Press.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*, 1–40. http://dx.doi.org/10.1023/A:1021302408382.

Calsyn, R. J., & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? *Journal of Educational Psychology, 69*, 136–145. http://dx.doi.org/10.1037/0022-0663.69.2.136.

Chen, J. A., Tutwiler, M. S., Metcalf, S. J., Kamarainen, A., Grotzer, T., & Dede, C. (2016). A multi-user virtual environment to support students' self-efficacy and interest in science: A latent growth model analysis. *Learning and Instruction, 41*, 11–22. http://dx.doi.org/10.1016/j.learninstruc.2015.09.007.

Chen, S.-K., Yeh, Y.-C., Hwang, F.-M., & Lin, S. S. J. (2013). The relationship between academic self-concept and achievement: A multicohort–multioccasion study. *Learning and Individual Differences, 23*, 172–178. http://dx.doi.org/10.1016/j.lindif.2012.07.021.

Demetriou, A., Platsidou, M., Efklides, A., Metallidou, Y., & Shayer, M. (1991). The development of quantitative-relational abilities from childhood to adolescence: Structure, scaling, and individual differences. *Learning and Instruction, 1*, 19–43. http://dx.doi.org/10.1016/0959-4752(91)90017-3.

Dong, Y., & Peng, C. Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus, 2*, 222. http://dx.doi.org/10.1186/2193-1801-2-222.

Durik, A. M., & Harackiewicz, J. M. (2007). Different strokes for different folks: How individual interest moderates the effects of situational factors on task interest. *Journal of Educational Psychology, 99*, 597–610. http://dx.doi.org/10.1037/0022-0663.99.3.5.

Durik, A. M., & Matarazzo, K. L. (2009). Revved up or turned off? How domain knowledge changes the relationship between perceived task complexity and task interest. *Learning and Individual Differences, 19*, 155–159. http://dx.doi.org/10.1016/j.lindif.2008.08.00.

Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., et al. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.). *Achievement and achievement motives: Psychological and sociological approaches* (pp. 75–146). San Francisco, CA: W. H. Freeman.

Eccles, J. S., & Wigfield, A. (1993). Age and gender differences in children's self- and task perceptions during elementary school. *Child Development, 64*, 830–847. http://dx.doi.org/10.1111/1467-8624.ep9308115032.

Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences, 19*, 499–505. http://dx.doi.org/10.1016/j.lindif.2009.05.004.

Finnish National Agency for Education (2017). *Finnish education in a nutshell.* Retrieved from http://www.oph.fi/download/146428_Finnish_Education_in_a_Nutshell.pdf.

Fredricks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology, 38*, 519–533. http://dx.doi.org/10.1037/0012-1649.38.4.519.

Fryer, L. K., Ainley, M., & Thompson, A. (2016). Modelling the links between students' interest in a domain, the tasks they experience and their interest in a course: Isn't interest what university is all about? *Learning and Individual Differences, 50*, 157–165. http://dx.doi.org/10.1016/j.lindif.2016.08.011.

Galyon, C. E., Blondin, C. A., Yaw, J. S., Nalls, M. L., & Williams, R. L. (2012). The relationship of academic self-efficacy to class participation and exam performance. *Social Psychology of Education, 15*, 233–249. http://dx.doi.org/10.1007/s11218-011-9175-x.

Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences, 47*, 182–193. http://dx.doi.org/10.1016/j.lindif.2016.01.002.

Gogol, K., Brunner, M., Goetz, T., Martin, R., Ugen, S., Keller, U., et al. (2014). "My questionnaire is too long!" the assessments of motivational-affective constructs with three-item and single-item measures. *Contemporary Educational Psychology, 39*, 188–205. http://dx.doi.org/10.1016/j.cedpsych.2014.04.002.

Graham, J., Tisher, R., Ainley, M., & Kennedy, G. (2008). Staying with the text: The contribution of gender, achievement orientations, and interest to students' performance on a literacy task. *Educational Psychology, 28*, 757–776. http://dx.doi.org/10.1080/01443410802260988.

Guo, J., Parker, P. D., Marsh, H. W., & Morin, A. J. S. (2015). Achievement, motivation, and educational choices: A longitudinal study of expectancy and value using a multiplicative perspective. *Developmental Psychology, 51*, 1163–1176. http://dx.doi.org/10.1037/a0039440.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94*, 562–575. http://dx.doi.org/10.1037/0022-0663.94.3.562.

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review, 13*, 191–209. http://dx.doi.org/10.1023/A:1016667621114.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*, 111–127. http://dx.doi.org/10.1207/s15326985ep4102_4.

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review, 17*, 63–84. http://dx.doi.org/10.1016/j.edurev.2015.11.002.

Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education, 28*, 1–35. http://dx.doi.org/10.1007/s10212-011-0097-y.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*, 1–55. http://dx.doi.org/10.1080/10705519909540118.

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*, 509–527. http://dx.doi.org/10.1111/1467-8624.00421.

Jansen, M., Lüdtke, O., & Schroeders, U. (2016). Evidence for a positive relation between interest and achievement: Examining between-person and within-person variation in five domains. *Contemporary Educational Psychology, 46*, 116–127. http://dx.doi.org/10.1016/j.cedpsych.2016.05.004.

Knogler, M., Harackiewicz, J. M., Gegenfurtner, A., & Lewalter, D. (2015). How situational is situational interest? Investigating the longitudinal structure of situational interest. *Contemporary Educational Psychology, 43*, 39–50. http://dx.doi.org/10.1016/j.cedpsych.2015.08.004.

Kriegbaum, K., Jansen, M., & Spinath, B. (2015). Motivation: A predictor of PISA's mathematical competence beyond intelligence and prior test achievement. *Learning and Individual Differences, 43*, 140–148. http://dx.doi.org/10.1016/j.lindif.2015.08.026.

Lane, A. M., Hall, R., & Lane, J. (2004). Self-efficacy and statistics performance among Sport Studies students. *Teaching in Higher Education, 9*, 435–448. http://dx.doi.org/10.1080/1356251042000252372.

Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences, 19*, 355–365. http://dx.doi.org/10.1016/j.lindif.2008.10.009.

Linnenbrink-Garcia, L., Durik, A. M., Conley, A. M., Barron, K. E., Tauer, J. M., Karabenick, S. A., et al. (2010). Measuring situational interest in academic domains. *Educational and Psychological Measurement, 70*, 647–671. http://dx.doi.org/10.1177/0013164409355699.

Linnenbrink-Garcia, L., Patall, E. A., & Messersmith, E. E. (2013). Antecedents and consequences of situational interest. *British Journal of Educational Psychology, 83*, 591–614. http://dx.doi.org/10.1111/j.2044-8279.2012.02080.x.

Liou, P.-Y. (2017). Profiles of adolescents' motivational beliefs in science learning and science achievement in 26 countries: Results from TIMSS 2011 data. *International Journal of Educational Research, 81*, 83–96. http://dx.doi.org/10.1016/j.ijer.2016.11.006.

Logtenberg, A., van Boxtel, C., & van Hout-Wolters, B. (2011). Stimulating situational interest and student questioning through three types of historical introductory texts. *European Journal of Psychology of Education, 26*, 179–198. http://dx.doi.org/10.1007/s10212-010-0041-6.

Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*, 59–77. http://dx.doi.org/10.1348/000709910X503501.

Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*, 397–416. http://dx.doi.org/10.1111/j.1467-8624.2005.00853.x.

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stilwell, C. D. (1989). Evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 105*, 430–445.

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology, 25*, 267–316.

Niemivirta, M., & Tapola, A. (2007). Self-efficacy, interest, and task performance: Within-task changes, mutual relationships, and predictive effects. *Zeitschrift für Pädagogische Psychologie, 21*, 241–250. http://dx.doi.org/10.1024/1010-0652.21.3.241.

OECD (2016). *PISA 2015 results (Volume I): Excellence and equity in education.* The Organisation for Economic Co-operation and DevelopmentParis: OECD Publishinghttp://dx.doi.org/10.1787/9789264266490-en.

Palmer, D. H. (2009). Student interest generated during an inquiry skills lesson. *Journal of Research in Science Teaching, 46*, 147–165. http://dx.doi.org/10.1002/tea.20263.

Patall, E. A., Vasquez, A. C., Steingut, R. R., Trimble, S. S., & Pituch, K. A. (2016). Daily interest, engagement, and autonomy support in the high school science classroom. *Contemporary Educational Psychology, 46*, 180–194. http://dx.doi.org/10.1016/j.cedpsych.2016.06.002.

Phan, H. P. (2014). Expectancy-value and cognitive process outcomes in mathematics learning: A structural equation analysis. *Higher Education Research and Development, 33*, 325–340. http://dx.doi.org/10.1080/07294360.2013.832161.

Renninger, K. A. (2009). Interest and identity development in instruction: An inductive model. *Educational Psychologist, 44*, 105–118. http://dx.doi.org/10.1080/00461520902832392.

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist, 46*, 168–184. http://dx.doi.org/10.1080/00461520.2011.587723.

Renninger, K. A., & Hidi, S. (2016). *The power of interest for motivation and learning.* New York/London: Routledge.

Rotgans, J. I., & Schmidt, H. G. (2011). Situational interest and academic achievement in the active-learning classroom. *Learning and Instruction, 21*, 58–67. http://dx.doi.org/10.1016/j.learninstruc.2009.11.001.

Rotgans, J. I., & Schmidt, H. G. (2014). Situational interest and learning: Thirst for knowledge. *Learning and Instruction, 32*, 37–50. http://dx.doi.org/10.1016/j.learninstruc.2014.01.002.

Schiefele, U. (2009). Situational and individual interest. In K. R. Wentzel, & A. Wigfield (Eds.). *Handbook of motivation at school* (pp. 197–222). New York/London: Routledge.

Schoor, C. (2016). Utility of reading — predictor of reading achievement? *Learning and Individual Differences, 45*, 151–158. http://dx.doi.org/10.1016/j.lindif.2015.11.024.

Schunk, D. H., & DiBenedetto, M. K. (2016). Self-efficacy theory in education. In K. R. Wentzel, & D. B. Miele (Eds.). *Handbook of motivation at school* (pp. 35–54). New York: Routledge.

Seaton, M., Marsh, H. W., Parker, P. D., Craven, R. G., & Yeung, A. S. (2015). The reciprocal effects model revisited: Extending its reach to gifted students attending academically selective schools. *Gifted Child Quarterly, 59*, 143–156. http://dx.doi.org/10.1177/0016986215583870.

Seegers, G., van Putten, C. M., & de Brabander, C. J. (2002). Goal orientation, perceived task outcome and task demands in mathematics tasks: Effects on students' attitude in actual task settings. *British Journal of Educational Psychology, 72*, 365–384. http://dx.doi.org/10.1348/000709902320634366.

Shen, B., Chen, A., & Guan, J. (2007). Using achievement goals and interest to predict learning in physical education. *Journal of Experimental Education, 75*, 89–108. http://dx.doi.org/10.3200/JEXE.75.2.89-108.

Silvia, P. J. (2003). Self-efficacy and interest: Experimental studies of optimal incompetence. *Journal of Vocational Behavior, 62*, 237–249. http://dx.doi.org/10.1016/S0001-8791(02)00013-1.

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology, 42*, 70–83. http://dx.doi.org/10.1037/0012-1649.42.1.70.

Skaalvik, E. M., & Hagtvet, K. A. (1990). Academic achievement and self-concept: An analysis of causal predominance in a developmental perspective. *Journal of Personality and Social Psychology, 58*, 292–307. http://dx.doi.org/10.1037/0022-3514.58.2.292.

Spinath, B., Eckert, C., & Steinmayr, R. (2014). Gender differences in school success: What are the roles of students' intelligence, personality and motivation? *Educational Research, 56*, 230–243. http://dx.doi.org/10.1080/00131881.2014.898917.

Spinath, B., & Steinmayr, R. (2008). Longitudinal analysis of intrinsic motivation and competence beliefs: Is there a relation over time? *Child Development, 79*, 1555–1569. http://dx.doi.org/10.1111/j.1467-8624.2008.01205.x.

Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research, 25*, 173–180. http://dx.doi.org/10.1207/s15327906mbr2502_4.

Sternberg, R. J., Castejón, J. L., Prieto, M. D., Hautamäki, J., & Grigorenko, E. L. (2001). Confirmatory factor analysis of the Sternberg Triarchic Abilities Test in three international samples: An empirical test of the triarchic theory of intelligence. *European Journal of Psychological Assessment, 17*, 481–496. http://dx.doi.org/10.1027//1015-5759.17.1.1.

Tanaka, A., & Murayama, K. (2014). Within-person analyses of situational interest and boredom: Interactions between task-specific perceptions and achievement goals. *Journal of Educational Psychology, 106*, 1122–1134. http://dx.doi.org/10.1037/a0036659.

Thoman, D. B., Sansone, C., & Pasupathi, M. (2007). Talking about interest: Exploring the role of social interaction for regulating motivation and the interest experience. *Journal of Happiness Studies, 8*, 335–370. http://dx.doi.org/10.1007/s10902-006-9016-3.

Vainikainen, M.-P. (2014). *Finnish primary school pupils' performance in learning to learn assessments: A longitudinal perspective on educational equity*Department of Education, Research Reports 360. University of Helsinki. Helsinki: Unigrafia.

Vainikainen, M.-P., Salmi, H., & Thuneberg, H. (2015). Situational interest and learning in a science center mathematics exhibition. *Journal of Research in STEM Education, 1*, 15–29.

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. http://dx.doi.org/10.1080/00313831.2014.904419.

Wang, Z., & Adesope, O. (2016). Exploring the effects of seductive details with the 4-phase model of interest. *Learning and Motivation, 55*, 65–77. http://dx.doi.org/10.1016/j.lmot.2016.06.003.

Wechsler, D. (1981). *WAIS-R manual: Wechsler adult intelligence scale-revised.* Psychological Corporation.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology, 25*, 68–81. http://dx.doi.org/10.1006/ceps.1999.1015.

Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbreton, A. J. A., Freedman-Doan, C., et al. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology, 89*, 451–469. http://dx.doi.org/10.1037/0022-0663.89.3.451.

Williams, T., & Williams, K. (2010). Self-efficacy and performance in mathematics: Reciprocal determinism in 33 nations. *Journal of Educational Psychology, 102*, 453–466. http://dx.doi.org/10.1037/a0017271.

Zhu, X., Chen, A., Ennis, C., Sun, H., Hopple, C., Bonello, M., et al. (2009). Situational interest, cognitive engagement, and achievement in physical education. *Contemporary Educational Psychology, 34*, 221–229. http://dx.doi.org/10.1016/j.cedpsych.2009.05.002.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*, 82–91. http://dx.doi.org/10.1006/ceps.1999.1016.