Does Task-specific Self-efficacy Predict Science Competencies?

Kristin Nyberg[1], Susanne Koerber[1], Christopher Osterhaus[2]

[1]Institute of Psychology, University of Education, Freiburg, Germany

[2]Institute of Psychology, University of Vechta, Vechta, Germany

Correspondence: Kristin Nyberg, Institute of Psychology, University of Education Freiburg, Kunzenweg 21, 79117 Freiburg, Germany.

Abstract

Self-efficacy is an affective-motivational factor that strongly predicts academic performance. With respect to science competencies, self-efficacy is related to two subcomponents that are closely associated already in kindergarten: Science content knowledge (e.g., physics knowledge) and scientific reasoning (e.g., knowing how to conduct a controlled experiment). To make accurate action predictions, the precise and specific measurement of self-efficacy is needed. With respect to different subcomponents of science competencies (i.e., science knowledge and scientific reasoning), there is to date a lack of studies that simultaneously investigate the association between students' self-efficacy and their performance in these two subcomponents of science competencies. The complex (cross-)relations between these constructs are investigated in the present study. The sample comprised N=181 fifth graders (90 girls, 91 boys). Exploratory and confirmatory factor analyses suggest that the two task-specific self-efficacy scales (scientific reasoning and science content knowledge) can be distinguished from each other and from general academic self-efficacy. Structural equation models reveal that task-specific self-efficacy in scientific reasoning is related to performance in scientific reasoning (.52) and science content knowledge (.32). Conversely, task-specific self-efficacy in science content knowledge correlates with performance in science content knowledge (.36) and scientific reasoning (.27). As expected, the strongest correlations between task-specific self-efficacy and performance emerge within the domain, but the significant cross-relations show the potential for furthering both aspects of performance and self-efficacy of science competencies and a need for a more detailed (longitudinal) investigation of these complex relations.

Keywords: self-efficacy, science competencies, scientific reasoning, science content knowledge, modelling self- efficacy, structural equation modeling

1. Introduction

There is substantial interest in educational research in the factors that can explain individual differences between students' academic performance (Elliott et al., 2019; Lee & Stankov, 2018). In educational developmental research on the influences of students' academic performance, motivational-affective factors play an increasingly important role (e.g., Lavrijsen et al., 2021). The results of a longitudinal study by Boncquet et al. (2020), which surveyed sixth graders in Belgium as they transitioned to secondary school, showed that motivational constructs explain substantial amounts of variation in mathematics achievement - even when controlling for intelligence. Among these affective-motivational variables, self-efficacy is an especially important factor that affects academic performance. A meta-analysis by Richardson et al. (2012) showed that study-related self-efficacy was the strongest predictor of study achievement among many motivational-affective variables (r = .59). Positive self-efficacy is however not only associated with academic performance; it is also related to motivational processes, self-regulation, self-perception and interest (Bandura & Schunk, 1981; Klassen & Usher, 2010; Pajares & Valiante, 1997; Schunk, 1995). Students with positive self-efficacy invest more persistence and effort in accomplishing challenging tasks (Pajares, 1996). Based on Bandura's social cognition theory, self-efficacy is classified as a competence belief (e.g., Bandura, 1997) and are understood as the belief in one's own ability to cope with a future situation or task.

According to Bandura (1986), self-efficacy should be viewed multidimensionally and not as a single disposition. Further, Bandura (1997) recommended that items should be used that measure the expectations and performance as close as possible to the identical tasks or situation. This recommendation presupposes that the more precisely self-efficacy is assessed, the more accurate the obtained action predictions will be (Bandura, 1997; Schwarzer & Jerusalem, 2002).

Consequently, self-efficacy is assumed to differ by domain or even task. That is, a student might have different levels of self-efficacy in language skills or mathematics or even a specific area of mathematics. Scales measuring self-efficacy exist for single subjects such as mathematics (Parker et al., 2014; Usher & Pajares, 2009), writing skills (Pajares & Valiante, 1997), or natural sciences such as biology (Baldwin et al., 1999) or physics (Gurcay & Ferah, 2018).

Domains such as mathematics, biology, and physics, are included in the curriculum as independent school subjects. Less is known about self-efficacy in more general domains like scientific reasoning (e.g., experimentation), which is a required topic of teaching in the elementary school curriculum in Germany (Ministerium für Kultus, Jugend und Sport, 2016).

Scientific reasoning can be viewed as a comprehensive construct that subsumes different skills, such as experimentation, understanding the nature of science, or data interpretation (Koerber et al., 2015). Studies show that these different competencies share a common conceptual core described mainly by an understanding of the relation between hypothesis and evidence (Kuhn, 2010; Osterhaus et al., 2017; Sodian, 2018). Validated group tests exist to reliably and comprehensively measure scientific reasoning in elementary school and above (e.g., SPR- I Osterhaus et al., 2020). Skills in scientific reasoning are also associated with important academic skills such as science content knowledge and further skills in science, technology, engineering, and mathematics (STEM) subjects (Labudde & Möller, 2012; Osborne, 2013; Ramseger, 2009; Schwichow et al.; 2020, Sodian & Koerber, 2007). Numerous studies have reported significant relations between scientific reasoning skills and science content knowledge in secondary school children (Songer & Linn, 1991; Stathopoulou & Vosniadou, 2007; Stender et al., 2018), elementary school (Pollmeier et al., 2017) and even at the end of kindergarten (Koerber & Osterhaus, 2019; 2021). Both constructs (scientific reasoning and science content knowledge) are also reflected in the conceptualization of scientific literacy in the PISA studies (OECD, 2007, 2016).

Several studies have examined the relation between self-efficacy and the performance in the natural sciences, including general chemistry (Ferrell et al., 2016; Ramnarain & Ramaila, 2018) and physics (Cavallo et al., 2004). Ardura and Galan (2019) administered the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1991) to 507 secondary school students to investigate, among other parameters, the relation between self-efficacy and performance in physics and chemistry (taught together in Spain in the last five years before entering university). The study found that self-efficacy and performance were related ($r = .31$).

Numerous studies have investigated in between scientific reasoning and more-general cognitive skills, such as intelligence, language skills, or executive functions (Koerber et al., 2015; Mayer et al., 2014; Osterhaus et al., 2017; van der Graaf et al., 2015). However, general cognitive skills alone cannot explain sufficient amounts of individual differences, leaving large portions of the variance in the performance of scientific reasoning skills unexplained. Few studies have investigated self-efficacy in science competencies like understanding science concepts, scientific literacy, or scientific reasoning, particularly in high school rather than college-age students. For example, a study by Liu et al. (2006) showed that the self-efficacy of sixth graders in science class significantly correlated with understanding science concepts ($r = .28$). Jansen et al. (2015) used a sample from the 2006 PISA survey in Germany to examine performance in scientific literacy and self-efficacy. Around 5,000 ninth graders were presented with a self-efficacy scale assessing the extent to which they thought they were confident in solving a task in scientific literacy (e.g., predict how changes in an environment will affect the survival of certain species). One finding of the study suggests self-efficacy is a significant predictor ($\beta = .55$) of performance in scientific literacy. The results of a study by Nyberg et al. (2022) found significant correlations between task-specific self-efficacy and scientific reasoning skills in Grades 4 ($r = .19$) and 8 ($r = .29$).

The reviewed studies have focused on the influence or relation between self-efficacy and single science competencies. Studies that simultaneously investigate the association between self-efficacy and students' performances in several subcomponents of science competencies, such as scientific reasoning skills and science content knowledge, have not yet been conducted.

Examining the structure of self-efficacy of different domains is an important step before investigating the relation between self-efficacy and the two science competencies (scientific reasoning and science content knowledge). Studies exist examining the structure of different self-efficacy facets (e.g., domain-specific vs. global). Lent et al. (1997) studied various self-beliefs in university students in the area of mathematics. They used self-efficacy scales that differed in the degree of domain specificity (specific mathematics content) and globality (broader academic skills). As criterion variables, overall grades were collected for domain specific performance in the area of mathematics and global outcomes. The reported factor analytic results revealed different latent dimensions, which supports the hypothesis that different self-efficacy facets can be separated. Lent et al. (1997) concluded that the different facets of self-efficacy seem to be useful in predicting domain-specific performance criteria.

Studies that investigate more than one (science) skill and measure task-specific self-efficacy with the same item pool (i.e., an item pool that is especially developed to measure the specific skill) seem to be lacking. Moreover, little is known about the structure and whether two task-specific self-efficacy scales of science competencies (scientific reasoning and science

content knowledge in the present study) are distinct dimensions. Similarly, few studies address (cross-)relations between different self-efficacy scales for different skill domains and the associated performance.

The present study assesses task-specific self-efficacy and the respective performance for two science constructs: Scientific reasoning and the conceptual understanding of science content knowledge. In a first step we (1) assess the reliability of the task-specific self-efficacy scale for scientific reasoning and science content knowledge. We then test (2) the multidimensional and context-specific structure of the self-efficacy scale for scientific reasoning and science content knowledge to examine whether the two task-specific self-efficacy dimensions are separate and distinct from more general academic self-efficacy. Finally, we analyze (3) (cross-)relations between the different self-efficacy scales and the performance measures in scientific reasoning and science content knowledge.

We expect that a task-specific scale would be reliable for the two domains, as demonstrated in other studies (e.g., Nyberg et al., 2022 for scientific reasoning skills and Siefer et al., 2020, for a specific area in mathematics). We expect that a differentiation of the three self-efficacy dimensions (scientific reasoning, science content knowledge and academic) will be successful, that is, a three-dimensional model will fit the data better than a one-factor model. This expectation is motivated by studies showing that a multidimensional structure of self-efficacy in other domains is highly context-dependent (Bandura, 1997; Miller et al., 1999; Skaalvik & Skaalvik, 2007; Zumbrunn et al., 2020), as well as by studies that show that self-efficacy dimensions of different specificity can be separated (e.g., Lent, 1997). The analysis of (cross-)relations between different self-efficacy scales and between the respective performances in different domains are under-researched. Based on studies that found correlations between skills in scientific reasoning and science content knowledge (Koerber & Osterhaus, 2019; Pollmeier et al., 2017), as well as findings that suggest close associations between self-efficacy in a special field and the corresponding performance, we predict significant relations.

We examined these questions in fifth graders because the transition to secondary education is an important stage in the development of self-efficacy. Elementary school students seem to overestimate their competencies, whereas with increasing age the self-evaluations are more and more in line with external evaluations (Nicholls, 1979; Stipek & Hofman, 1980; Pajares & Schunk, 2001).

The results of this study will provide important information about the structure and measurement of self-efficacy in scientific reasoning and science content knowledge and how they relate to each other. Therefore, these results could offer important implications for intervention, especially with regard to the important promotion of skills in the STEM subjects.

2. Method

2.1 Participants

A total of $N = 181$ fifth graders participated in the study (90 girls, 91 boys; $M_{age}=11.04$ years; SD=6 months). The students were recruited from 10 high and middle schools close to a mid-sized city in southern Germany. Of the 181 children, 32.6% spoke at least one language other than German at home (the most frequently reported languages were Russian, Kurdish, and English). Student assent and written consent from caretakers were obtained for all participants.

2.2 Study Materials

2.2.1 Performance Measures

Science content knowledge. Science content knowledge was assessed with a total of five tasks (Baumert et. al., 1999; Pollmeier et. al., 2017). The reliability of the entire scale was Cronbach's $\alpha = .58$. The scale measures science content knowledge, however, it captures several aspects (e.g., displacement or energy transformation). The reliability remains in a similar range compared to other studies that use a scale to measure multiple aspects. (e.g., Koerber et al., 2015 for another domain). The reported reliabilities of > 5 can be interpreted for scales especially with few items (e.g., Nunnally & Bernstein, 1978).

Three of the tasks (bvr1, bvr4, bvr6) were multiple-select tasks that captured students' understanding of floating and sinking, in particular density, displacement, and buoyancy (Table 1). The items were selected from an instrument by Pollmeier et al. (2017) and Pollmeier et al. (2011). For these tasks, three answer options were presented, and for each answer option, students indicated whether they accepted or rejected the option. At the end of each task, they specified which of the three options was the best answer. The tasks were coded dichotomously, that is, 1 point was given for a correct answer and 0 points for an incorrect answer to each of the three forced-choice items. Two further tasks (Item numbers: S042404 (glass pitcher, b13) and S042407 (ice blocks, b20) were taken from the released International Association for the Evaluation of Educational Achievement (2013) items from the Trends in International Mathematics and Science Study (TIMSS) in 2011. The tasks measured students' understanding of physical concepts of changes in matter and energy transformation (Table 1). The two tasks had an open answer format and were coded dichotomously. For the correct answer 1 point was given and 0 points for an incorrect answer.

Scientific reasoning. To measure performance in scientific reasoning, the SPR-I(7) was used. The SPR-I(7) contains seven items from the Science-P Reasoning Inventory (SPR-I; Koerber et al., 2015; Osterhaus et al., 2020) and captures broad scientific reasoning skills (three items on the nature of science [NOS], three items on experimentation [EXP] and one item on data interpretation [DAT; Table 1]).Studies indicate that the sub-aspects of scientific reasoning skills share a common conceptual core (e.g. Koerber et. al, 2015). The SPR-I(7) has been validated and is particularly suitable for economical use (Osterhaus et al., 2020). The reliability of the here used scale was Cronbach's $\alpha = .51$ and show similar reliabilities as the study by Osterhaus et al. (2020). The coding process follows the SPR-I. For each task, three answer options were given (a naïve level = 0 points, an intermediate level = 1 point, a scientifically advanced level = 2 points) and for each answer option, students indicated whether they accept or reject the option. The lowest answer level selected was taken as the final score on the entire item (see Koerber et al., 2015, and Osterhaus et al., 2020, for further coding details).

Table 1. Overview of used items

Item Number	Skill component	Coding
SPR-I(7) (Osterhaus et al., 2020)		
A08	EXP (testing hypo. vs. prod. effects)	Naive (0), intermediate (1), advanced (2)
A10	EXP (testing hypo. vs. prod. effects)	Naive (0), intermediate (1), advanced (2)
A02	EXP (CVS)	Naive (0), intermediate (1), advanced (2)
A06	DAT (confounded data)	Naive (0), intermediate (1), advanced (2)
A05	NOS (NOS concepts)	Naive (0), intermediate (1), advanced (2)
A03	NOS (framework theories)	Naive (0), intermediate (1), advanced (2)
A11	NOS (NOS concepts)	Naive (0), intermediate (1), advanced (2)
SCK (Pollmeier et al., 2017)		
Bvr1	Density, displacement, and buoyancy	Best answer incorrect (0), correct (1)
Bvr4	Density, displacement, and buoyancy	Best answer incorrect (0), correct (1)
Bvr6	Density, displacement, and buoyancy	Best answer incorrect (0), correct (1)
TIMSS (Baumert et al., 1999)		
B13	Evaporation	Incorrect answer (0), correct answer (1)
B20	Evaporation	Incorrect answer (0), correct answer (1)

SPR-I = Science-P Reasoning Inventory, EXP = Experimentation, hypo. = hypotheses, prod.= producing, CVS = Control of Variable Strategy, DAT = Data interpretation, NOS = Nature of Science, SCK= Science content knowledge

2.2.2 Self-efficacy Scales

Academic self-efficacy. Academic self-efficacy was measured using five items (Item no. 1, 2, 4, 6, 7) of the scale (WIRKSCHUL) by Jerusalem and Satow (1999), for example, "I can solve even the difficult tasks in class if I exert myself (own translation)". The items were assessed using a 4-point Likert scale ranging from "strongly disagree" (0) to "strongly agree" (3). The scale was applied in German.

Task-specific self-efficacy. Task-specific self-efficacy was measured with three 4-point Likert scaled items, presented before each of the five science content knowledge and the seven scientific reasoning tasks. Before each task on scientific reasoning and science content knowledge, the task-specific self-efficacy was measured with the following three items: 1) "I know how to deal with the task" 2) "I am very familiar with such tasks and know how to solve it" 3) "I would need help to solve the task." The Likert scale ranged from "strongly disagree" (0) to "strongly agree" (3). For the factor analytic calculations and structural equation models, the three items per task were combined into one item value, resulting in five task-specific self-efficacy values for science content knowledge and seven for scientific reasoning.

2.3 Study Procedure

The testing took place in the participating schools as a whole-class testing procedure. The booklet consisted of illustrated paper-pencil items. Students worked individually in their booklet. To avoid confounding effects of reading ability, the items were presented by a PowerPoint presentation and read aloud by an experimenter. Test assistants supported the students with the procedure and ensured that students worked in their own booklet. Before each performance task (scientific reasoning and science content knowledge) students completed the task-specific self-efficacy items. The instruction was to look at the performance task for 25 seconds but to not try to solve the task. After, they filled in the task-specific self-efficacy, they solved the respective performance task. The testing took about 60-70 min.

3. Results

3.1 Core Performance and Reliability of the Task-specific Self-efficacy Scales

Figure 1 shows the mean percent of performance in scientific reasoning and science content knowledge and task-specific

self-efficacy in scientific reasoning and science content knowledge plus the academic self-efficacy transformed to a percentage scale. The percentages represent the mean points achieved on the respective scale.

Performance measures. In the scientific reasoning tasks, the fifth graders scored on average 6.75 (SD = 2.58) out of 14 points (48.2%) compared to an average of 2.30 (SD = 1.39) points out of a potential 5 points (46.1%) for the performance in science content knowledge.

Self-efficacy scales. The mean score achieved in task-specific self-efficacy in scientific reasoning was 37.37 (SD = 9.04) out of 63 points (59.3%). The internal consistency was high (Cronbach's α = .85). The corrected item-total correlations (r_{it} = .48-.61) were high for most of the items. The item characteristics hence indicate good reliability of the scale (Bortz & Döring, 2006; Moosbrugger & Kleava, 2012). For task-specific self-efficacy of science content knowledge, the mean score was 30.59 (SD = 8.60) out of 45 points (67.98%). The internal consistency was high (Cronbach's α = .90) and so were the corrected item-total correlations (r_{it} = .52-.70) which show a good reliability of the scale (Bortz & Döring, 2006; Moosbrugger & Kleava, 2012). In academic self-efficacy, the students stated an average of 11.16 (SD = 2.27) of 15 points (74.4%), with Cronbach's α = .60.
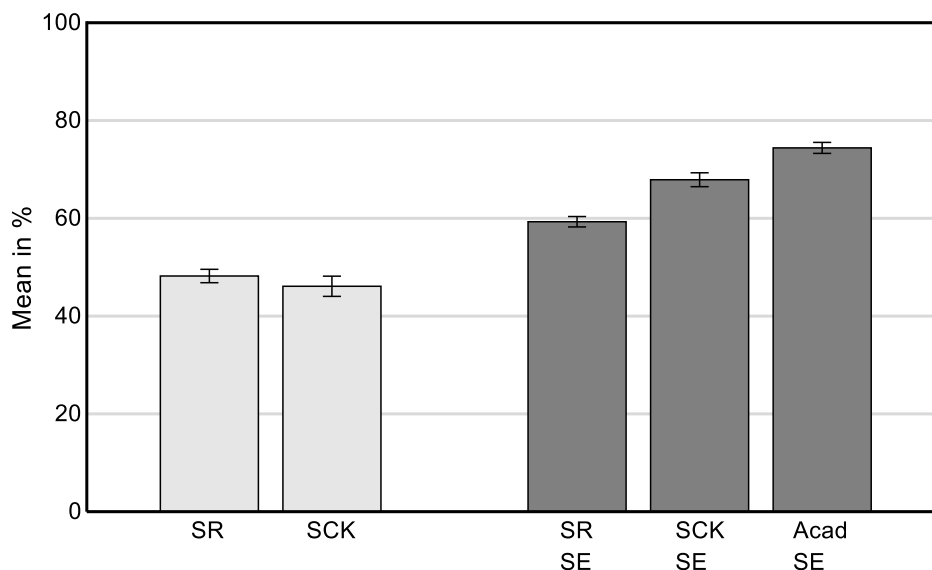

Figure 1. Core performance in the performances and self-efficacy scales, SR= scientific reasoning, SCK= science content knowledge, SE= self-efficacy, Acad= academic

3.2 Analysis of the Self-efficacy Scales

3.2.1 Exploratory Factor Analysis

To test whether the three self-efficacy scales (task-specific self-efficacy in scientific reasoning, task-specific self-efficacy in science content knowledge and academic self-efficacy) are conceptually separate self-efficacy scales, an exploratory factor analysis (EFA) was conducted using principal component analysis and varimax rotation. Eigenvalues greater than 1 were included. The results indicated that 13 of the 17 items could be discriminately assigned to the three factors scientific reasoning, science content knowledge, and academic self-efficacy. All factor loadings of the 13 items were between .51 and .89. The remaining four items could not clearly be assigned to any of the three factors, but rather formed two additional factors with two variables each, which were difficult to interpret both statistically (cross-loadings, only two items on one factor) and in terms of content. Therefore, these four items were eliminated for the subsequent analysis.

Table 2. Factor loading of the items of the three self-efficacy scales

SE Item	Factor		
	SR	SCK	Acad
A11	**0.799**		
A03	**0.777**		
A06	**0.775**		
A08	**0.522**		
A05	**0.518**		
A10	**0.509**		
Bvr4		**0.889**	
Bvr1		**0.816**	
Bvr6		**0.750**	
B13		**0.545**	
Acad 1			**0.720**
Acad 6			**0.677**
Acad 4			**0.573**

SE= self-efficacy, SR= scientific reasoning, SCK= science content knowledge, Acad= academic

3.2.2 Confirmatory Factor Analysis

To confirm the three-dimensional structure of the self-efficacy scales, a confirmatory factor analysis (CFA) with the remaining 13 items was performed. For the analysis, the lavaan package within the statistical software R was applied (Rosseel, 2012). The lavaan package is fully open-source and enables latent variable modeling, including confirmatory factor analysis and structural equation modeling.

The three-dimensional (task-specific self-efficacy in scientific reasoning, science content knowledge, and academic self-efficacy) model fitted the data well (comparative fit index [CFI] = .95, Tucker-Lewis index [TLI] = .93, and root-mean-square error of approximation [RMSEA] = .06 [.03-.08], Model 1 in Table 3). Factor loadings were high, for all items > .5 and many even higher (Table 2).

To test whether the 3-factor structure of the measured self-efficacy scales reflected our data better than other models, different models were compared. To test the multidimensionality of the measured self-efficacy scales, all items were loaded on a single factor in Model 2. Model 3 tested a two-factor structure, in which the task-specific scientific reasoning and science content knowledge scales were merged as one factor and academic self-efficacy as a separate factor. The results of the CFA revealed poor fit of the data for Model 2 and 3 (Table 3).

Model 3a and b examined whether task-specific self-efficacy in scientific reasoning and science content knowledge can be differentiated from each other when the academic self-efficacy scale is excluded from the model. Model 3a tested therefore a two-factor structure with task-specific self-efficacy in scientific reasoning and science content knowledge and 3b tested a one-factor structure with self-efficacy in scientific reasoning and science content knowledge as one factor. Model 3a showed acceptable fit indices to the model, whereas model 3b showed poor fit (Table 3).

The results of the CFA indicate the three-factor model with task-specific self-efficacy in scientific reasoning and science content knowledge and academic self-efficacy (Model 1) to have the best model fit. This result suggests, as hypothesized earlier, that self-efficacy is multidimensional and context specific.

Table 3. Fit-statistics of the different CFA models

							Comparison		
Model	χ^2	df	TLI	CFI	RMSEA	SRMR	$\Delta\chi^2$	Δdf	p
Model 1 3-Factors (SR, SCK and acad)	92.01	62	.93	.95	0.063	.0067			
Model 2 1-Factor (SR, SCK and acad)	184.82	65	.73	.78	0.123	.081	92,81[1]	3	.001
Model 3 2-Factors (SR/SCK and acad)	171.10	64	.76	.80	0.118	.079	79.09[1]	2	.001
Model 3a 2-Factors (SR and SCK)	66.24	56	.91	.94	0.089	.069			
Model 3b 1-Factor (SR/SCK)	147.81	58	.70	.77	0.163	.090	81,57[2]	2	.001

SR = scientific reasoning, SCK = science content knowledge, acad = academic, SC/SCK = SR and SCK merged as one factor. [1]Comparison Model 1. [2]Comparison Model 3a

3.2.3 Relations Between Performance Measures and Self-efficacy Scales

To analyze the relations between the three self-efficacy scales (scientific reasoning, science content knowledge and academic) and the performances, structural equation models (SEM) were computed using the lavaan package for R (Rosseel, 2012). Various theoretical models were tested to investigate the (cross-)relations between the specific (scientific reasoning and science content knowledge) and more general measures of self-efficacy (academic) scales and the respective performance. In the analyses, self-efficacy scales were included as latent variables and the two performance measures as manifest variables (Figure 2).

Model 2 (Table 4) included the two task-specific self-efficacy scales and the two performance measures, but not academic self-efficacy. This model showed an excellent and thus best fit of the data to one of the postulated models (CFI= .95, TLI = .93, and RMSEA= .067 [.03-.11]; Hu & Bentler, 1999; Kline, 2015).

In Model 2, the strongest association emerged between task-specific self-efficacy in scientific reasoning and the performance in scientific reasoning (.52). The scientific reasoning performance was also related to students' science content knowledge (.24) and their task-specific self-efficacy in science content knowledge (.32). Performance in science content knowledge was less related to task-specific self-efficacy for science content knowledge (.36) than the relation between task-specific self-efficacy for scientific reasoning and the particular performance. Performance in science content knowledge was also related to task-specific self-efficacy in scientific reasoning (.27). Additionally, the two task-specific self-efficacy scales were most strongly related (.56).

Model 1, which additionally included the academic self-efficacy scale, showed poorer fit indices (CFI= .92, TLI = .90, and RMSEA= .075) than Model 2. The paths between the variables revealed that academic self-efficacy was related to task-specific self-efficacy in scientific reasoning (.45) and science content knowledge (.64). Academic self-efficacy was not significantly related to the performance in scientific reasoning (.08) or science content knowledge (.06).

Models 2a-c analyzed relations between the variables by setting certain paths to 0. In Model 2a, the direct paths between task-specific self-efficacy and performance in science content knowledge and scientific reasoning (but not the cross-relations) were set to 0. The models indicated poorer fit of the data (CFI= .91, TLI = .89, and RMSEA= .082 [.04-.12]), than Model 2. In Model 2b and 2c, the cross-relations between task-specific self-efficacy and performance in scientific reasoning and science content knowledge were constrained to be 0. That is, in Model 2b, the path from task-specific self-efficacy in science content knowledge to performance in scientific reasoning was set to 0; in Model 2c, the path from task-specific self-efficacy of scientific reasoning to performance in science content knowledge was determined to be 0. Both models (2b and c) revealed a poorer model fit than Model 2 (Table 4), thus confirming the cross-relations between task-specific self-efficacy in scientific reasoning, science content knowledge and the respective performance (Figure 2).

Table 4. Fit statistics of the different SEM models

Model	$x2$	df	CFI	TLI	RMSEA	SRMR
Model 1 (SE in SR, SCK and acad + performance in SR and SCK)	113.91	82	.92	.90	.068	.075
Model 2 (SE in SR and SCK + performance in SR and SCK)	70.28	66	.95	.93	.067	.073
Model 2a (SE in SR to SR; SE in SCK to SCK)	82.25	52	.91	.89	.082	.125
Model 2b (SE in SR to SR and SCK; SE in SCK to SCK)	75.64	51	.93	.91	.075	.096
Model 2c (SE in SCK to SCK and SR; SE in SR to SR)	78,13	51	.93	.90	.077	.081

SE = self-efficacy, SR = scientific reasoning, SCK = science content knowledge, acad = academic
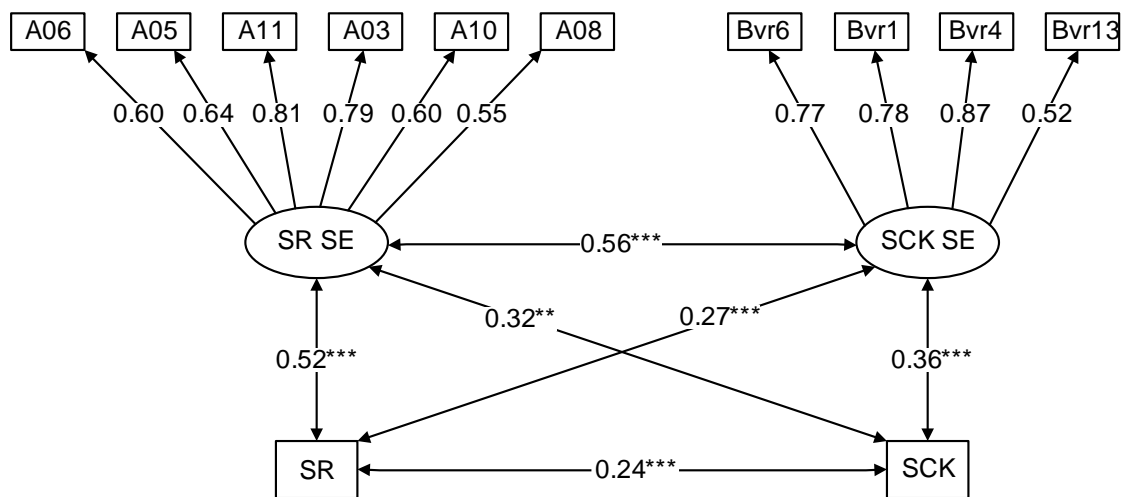
Figure 2. SEM Model with task-specific self-efficacy scales and performance measures, SR = scientific reasoning (performance), SCK = science content knowledge (performance), SE = self-efficacy, **p< .01, ***p< .001

4. Discussion

The main aims of the study were (1) to measure self-efficacy reliably and as precisely as possible (here task-specific) with the same item pool for different science competencies (scientific reasoning and science content knowledge) (2) to test the multidimensionality, context-specific structure of self-efficacy for the two constructs, scientific reasoning and scientific content knowledge, and to test whether the two task-specific dimensions can be differentiated from general academic self-efficacy, and finally, (3) to analyze (cross-)relations between the different self-efficacy scales and the performances in scientific reasoning and science content knowledge.

Overall, the findings show significant relations between task-specific self-efficacy in scientific reasoning and performance in scientific reasoning as well as between task-specific self-efficacy in science content knowledge and performance in science content knowledge. Also, our results reveal further cross-relations between the different constructs, suggesting

significant associations between all constructs investigated. In general, our results thus suggest a close association of the two concepts (scientific reasoning and science content knowledge); nonetheless, the task-specific self-efficacy scales emerged as independent dimensions for the two constructs, suggesting that a separate assessment of these subcomponents of general science self-efficacy is indicated.

4.1 Core Performance and Reliability of the Task-specific Self-efficacy Scales

A review of the literature revealed a research gap concerning studies measured self-efficacy for different skill domains with the same item pool with the purpose of comparing the domains and their relation to performance. Furthermore, studies suggest to measure self-efficacy as specifically and precisely as possible to generate better action predictions (Bandura, 1997; Schwarzer & Jerusalem, 2002). For this purpose, the present study measured task-specific self-efficacy for the two skill domains of scientific reasoning and science content knowledge.

The item characteristics (internal consistency and corrected item-total correlations) of the reliability analysis indicate that the items can be reliably used to measure self-efficacy in scientific reasoning and science content knowledge in a task-specific manner.

Looking at core performance, the students notably rated themselves as self-efficacious in all three scales and no ceiling effects occurred. The students reported the highest scores for academic self-efficacy, followed by self-efficacy in science content knowledge, and the lowest scores were reported for self-efficacy in scientific reasoning. These findings are consistent with the idea that mastery experiences are one of the most important sources for shaping and developing self-efficacy (Usher & Pajares, 2009; Nagengast et al., 2011). Students in the fifth grade had the opportunity to benefit from mastery experiences in the academic setting for many years, especially related to general academic skills in language or mathematics. Students also have more frequent opportunities to benefit from mastery experiences in the area of science content knowledge relative to scientific reasoning skills, which is because scientific reasoning skills are not so dominant in the curriculum. With the 2016 curriculum, science education explicitly requires that students in Grades 3 and 4 know specific experiments and have basic skills of experimentation. However, opportunities to explicitly work on such problems and receive informative feedback on it are less frequent than for general academic skills or science content knowledge.

The SPR-I(7) still sufficiently differentiates performance in fifth grade. In line with previous studies (e.g., Bullock et al., 2009), scientific reasoning skills appear to continue to develop from elementary to secondary school and may not even be fully developed by the end of secondary school (Bullock et al., 2009). This continued development again underscores the importance of studying the factors that relate to the development and learning of scientific reasoning skills. Given the high potential for development in scientific reasoning skills during the school years, especially in the context of the important STEM subjects, self-efficacy may be a valuable factor that can be used to develop interventions.

4.2 Multidimensionality and Context-specific Structure of the Different Self-efficacy Scales

Both the EFA and CFA results suggest a conceptual separation of task-specific self-efficacies (scientific reasoning and science content knowledge) and general academic self-efficacy. The model that included all three self-efficacy scales as separate factors in the CFA showed the best model fits relative to the poor fit indices of the comparison models. In the comparison models, for example, self-efficacy in scientific reasoning and science content knowledge were combined into one factor. The analysis demonstrated that the task-specific self-efficacy in scientific reasoning should be distinguished from task-specific self-efficacy in science content knowledge as well as from general academic self-efficacy. This separation of self-efficacy facets with different specificity is backed up by research. For example, Lent et al. (1997) argued for a separation of different facets of self-efficacy (specific mathematics content and broader academic skills). The multidimensionality of self-efficacy is also found in other studies (Skaalvik &Skaalvik, 2007). The results are consistent with Bandura's (1977; 1986) work, which suggests that self-efficacy should be viewed multidimensionally rather than as a single disposition.

The specific measurement and differentiation from an all-purpose measurement of self-efficacy can be useful for the further development and fostering of important science competencies. The successful mastery of STEM subjects requires both scientific reasoning skills and knowledge of science content. Thus, fostering the various task-specific skills is of high relevance. Promoting specificity could target the specific self-efficacy of the two skill domains. This task-specific focus of self-efficacy is also backed up by studies showing that skills related to scientific reasoning (e.g., understanding science concepts or scientific literacy) correlate positively with self-efficacy (Liu, 2006), and self-efficacy appears to be a significant predictor of performance in scientific literacy (Jansen, 2015). This relation or influence is meaningful because self-efficacy is a construct that can be improved through training (Margolis & McCabe, 2006). In the context of school, teachers' knowledge of the different forms of self-efficacy is crucial, and they should be aware that these different forms can be addressed to improve student skills.

4.3 Relations Between Different Self-efficacy Scales and Performances'

In self-efficacy research, studies have captured and analyzed self-efficacy scales for single school subjects or skill areas (e.g., mathematics or physical education). Little is known about studies that have investigated self-efficacy for more than one school subject or skill area with cross-relations. These cross-relations seem particularly meaningful for skill domains that are correlated and relevant to higher-order skills such as STEM subjects (in our study). Our results of the SEMs indicate the importance of the cross-relations between the self-efficacy scales and with performance in particular. The model that included the two specific self-efficacy scales (scientific reasoning and science content knowledge without academic self-efficacy) and the performances demonstrated the best model fit, revealing positive relations between the task-specific self-efficacy scales and the respective performances. Furthermore, a weaker relation between task-specific self-efficacy and the other performance was shown, that is, task-specific self-efficacy in scientific reasoning is related to performance in science content knowledge and task-specific self-efficacy in science content knowledge is related to performance in scientific reasoning. The SEM model in which academic self-efficacy was added showed poorer model fit and no significant relations to performance in scientific reasoning or science content knowledge. The models that excluded the cross-relations between the performances also showed poorer model fit. These results from the SEMs support that (1) a precise and specific measurement of self-efficacy shows the highest correlation to the respective performance and that (2) cross-relations between specific self-efficacy and performance are relevant and should be analyzed.

4.4 Limitations and Future Directions

The present study gives a general overview of the cross-sectional structure and relations between scientific reasoning, science content knowledge, and self-efficacy. However, only longitudinal studies can provide information about the development of the relation and whether the specific measurement of self-efficacy is still suitable in later school years, especially with regard to cross-relations. In general, the results related to the performance measures should be interpreted carefully, as the scales do not show very high reliability. Nevertheless, significant relations emerged between the constructs. This suggests that, if at all, our study may have underestimated the strength of the associations between the constructs investigated, and that stronger associations may when scales with higher internal consistency are used. Analyzing the extent that more experience in the specific skills influences the relation among the three constructs would be fruitful, as well as the influence on different skills of science competencies. Further research should also consider the distribution of the correlation between skills in science competencies and self-efficacy, given that studies (Nyberg et al., 2022; Siefer et al., 2020) have shown that the relation between task-specific self-efficacy and performance is not homogeneously distributed, that is, an increase in self-efficacy is not necessarily associated with higher performance. Nyberg et al. (2022) reported that the correlation between task-specific self-efficacy and performance in scientific reasoning was not equally distributed within the sample and across different age groups. Groups of students who (substantially) over- or underestimate themselves emerge. These results imply that not all students would benefit equally from self-efficacy training. Studies that analyze this correlation within a sample in a differentiated way would help to advance the research in this area. This research could inform educators about possible interventions with regard to the important STEM subjects.

4.5 Implications

Based on these findings, theoretical and practical implications can be drawn. From a theoretical perspective, it can be inferred that (1) as one of the first studies that distinguishes self-efficacy in science competencies at different levels using SEM analyses, general academic self-efficacy is not significantly related to performance in scientific reasoning and science content knowledge at the beginning of secondary school. This is a noteworthy finding given that other studies have often sampled older secondary school students or college students (2) Task-specific self-efficacy is most strongly related to the corresponding performance (scientific reasoning or science content knowledge) but also to the respective other performance. Nevertheless, the two task-specific self-efficacy scales form their own dimensions and underscore the differentiation of the two constructs. The practical implications relate mainly to teachers: (1) To implement influencing target promotion interventions for teachers, educators should know about the specificity of self-efficacy in science competencies, (2) know that specific self-efficacy has a relation to specific skills like scientific reasoning and can be fostered, and (3) be aware that specific self-efficacy have cross-relations to other skills, possibly mainly to skills that are correlated with each other (here scientific reasoning and science content knowledge) and thus promoting specificity can indirectly influence different skills.

Acknowledgement

References

Ardura, D., & Galán, A. (2019). The interplay of learning approaches and self-efficacy in secondary school students' academic achievement in science. *International Journal of Science Education, 41*, 1723-1743. https://doi.org/10.1080/09500693.2019.1638981

Baldwin, J. A., Ebert-May, D., & Burns, D. J. (1999). The development of a college biology self-efficacy instrument for nonmajors. *Science Education*, *83*, 397-408. https://doi.org/10.1002/(SICI)1098-237X(199907)83:4<397::AID-SCE1>3.0.CO;2-%23

Bandura, A. (1977). Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191-215.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Engelwood Cliffs.

Bandura, A. (1997). *Self efficacy: The exercise of control*. Freeman.

Bandura, A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41,* 586-598.

Baumert, J., Bos, W., Klieme, E., Lehmann, R., Lehrke, M., Hosenfeld, I., et al. (1999). *Testaufgaben zu TIMSS/III Mathematisch-naturwissenschaftliche Grundbildung und voruniversitäre Mathematik und Physik der Abschlußklassen der Sekundarstufe II (Population 3).* Max-Planck-Institut für Bildungsforschung.

Boncquet, M., Soenens, B., Verschueren, K., Lavrijsen, J., Flamant N., & Vansteenkiste, M. (2020). Killing two birds with one stone: The role of motivational resources in predicting changes in achievment and school well-being beyond intelligence. Contemporary *Educational Psychology, 63*, 101905. https://doi.org/10.1016/j.cedpsych.2020.101905

Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* [Research Methods and Evaluation for Social Scientists]. Springer.

Bullock, M., Sodian, B., & Koerber, S. (2009). Doing experiments and understanding science: Development of scientific reasoning from childhood to adulthood. In W. Schneider & M. Bullock (Eds.). *Human development from early childhood to early adulthood: Findings from a 20 year longitudinal study* (pp.173-197). Psychology Press.

Cavallo, A.M.L., Potter, W.H. & M. Rozman. 2004. Gender differences in learning constructs, shifts in learning constructs, and their relationship to course achievement in a structured inquiry, year long college physics course for life science majors. *School Science and Mathematics, 104*, 288-300. https://doi.org/10.1111/j.1949-8594.2004.tb18000.x

Elliott, J., Stankov, L., Lee, J., & Beckmann, J. F. (2019). What did PISA and TIMSS ever do for us? The potential of large scale datasets for understanding and improving educational practice. *Comparative Education, 55*, 133-155. https://doi.org/10.1080/03050068.2018.1545386

Ferrell, B., Phillips M. M. and Barbera J., (2016), Connecting achievement motivation to performance in general chemistry, *Chemistry Education Research and Practice., 17*, 1054-1066. https://doi.org/10.1039/C6RP00148C

Gurcay, D., & Ferah, H. O. (2018). High school students' critical thinking related to their metacognitive self-regulation and physics self-efficacy beliefs. *Journal of Education and Training Studies, 6*, 125-130. https://doi.org/10.11114/jets.v6i4.2980

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

Jansen, M., Scherer, R., & Schroeders, U. (2015). Students' self-concept and self-efficacy in the sciences: Differential relations to antecedents and educational outcomes. *Contemporary Educational Psychology, 41*, 13-24. https://doi.org/10.1016/j.cedpsych.2014.11.002

Jerusalem, M., & Satow, L. (1999). Schulbezogene Selbstwirksamkeitserwartung. In R. Schwarzer & M. Jerusalem (Eds.). *Skalen zur Erfassung von Lehrer- und Schülermerkmalen*. Freie Universität Berlin.

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.), *The Decade Ahead: Theoretical Perspectives on Motivation and Achievement* (pp. 1-33). Emerald Group Publishing.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilford.

Koerber, S., & Osterhaus, C. (2019). Individual differences in early scientific thinking: Assessment, cognitive influences, and their relevance for science. *Learning. Journal of Cognition and Development, 20*, 510-533. https://doi.org/10.1080/15248372.2019.1620232

Koerber, S., & Osterhaus, C. (2021). Science competencies in kindergarten: A prospective study in the last year of

kindergarten. *Unterrichtswissenschaft, 49*, 117-136. https://link.springer.com/article/10.1007/s42010-020-00093-55

Koerber, S., Mayer, D., Osterhaus, C., Schwippert, K., & Sodian, B. (2015). The Development of Scientific Thinking in Elementary School: A Comprehensive Inventory. *Child Development, 86*, 327-336. https://doi.org/10.1111/cdev.12298

Kuhn, D. (2010). What is scientific thinking and how does it develop? In U. Goswani (Ed.), *Handbook of childhood cognitive development* (2nd ed., pp. 497-523). Blackwell.

Labudde, P., & Möller, K. (2012). Stichwort: Naturwissenschaftlicher Unterricht. *Zeitschrift Für Erziehungswissenschaft, 15,* 11-36. https://doi.org/10.1007/s11618-012-0257-0

Lavrijsen, J., Vansteenkiste, M., Boncquet, M., & Verschueren, K. (2021). Does motivation predict changes in academic achievement beyond intelligence and personality? A multitheoretical perspective. *Journal of Educational Psychology, 114,* 772-790. https://doi.org/10.1037/edu0000666

Lee, J., & Stankov, L. (2018). Non-cognitive predictors of academic achievement: Evidence from TIMSS and PISA. *Learning and Individual Differences, 65*, 50-64. https://doi.org/10.1016/j.lindif.2018.05.009

Lent, R. W., Brown, S. D., & Gore, P. A. (1997). Discriminant and predictive validity of academic self-concept, academic self-efficacy, and mathematics-specific self-efficacy. *Journal of Counseling Psychology, 44*, 307-315. https://doi.org/10.1037/0022-0167.44.3.30

Liu, M., Hsieh, P., Cho, Y., & Schallert, D. L. (2006). Middle school students' self-efficacy, attitudes, and achievement in a computer-enhanced problem-based learning Environment. *Journal of Interactive Learning Research, 17*, 225-242.

Margolis, H., & McCabe, P. P. (2006). Improving self-efficacy and motivation: What to do, what to say. *Intervention In School & Clinic, 41*, 218-227. https://10.1177/10534512060410040401

Mayer, D., Sodian, B., Koerber, S., & Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities. *Learning and Instruction, 29*, 43-55. https://doi.org/10.1016/j.learninstruc.2013.07.005

Miller, J. W., Coombs, W. T., & Fuqua, D. R. (1999). An examination of psychometric properties of Bandura's Multidimensional Scales of Perceived Self-Efficacy. *Measurement and Evaluation in Counseling and Development, 31*, 186-196.

Ministerium für Kultus, Jugend und Sport (2016). *Bildungsplan der Grundschule für den Sachunterricht in Baden-Württemberg*.

Moosbrugger, H., & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion* [Test theory and questionnaire construction]. Springer.

Nagengast, B., Marsh, H. W., Scalas L.F., Xu, M. K., Hau, K. T., & Trautwein, U. (2011). Who took the "x" out of expectancy-value theory? a psychological mystery, a substantive-methodological synergy, and a cross-national generalization. *Psychological Science, 22*, 1058-1066. https://doi.org/10.1177/0956797611415540

Nicholls, J. G. (1979). Development of perception of own attainment and causal attributions for success and failure in reading. *Journal of Educational Psychology*, *71*, 94-99. https://doi.org/10.1037/0022-0663.71.1.94

Nunnally, J. C., & Bernstein, I. H. (1978). *Psychometric Theory*. McGraw-Hill

Nyberg, K., Koerber, S., & Osterhaus, C. (2022). Self-effective scientific reasoning? Differences between elementary and secondary school students. *Frontline Learning Research. 10,* 25-45. https://doi.org/10.14786/flr.v10i1.955

OECD. (2007). PISA 2006: Science competencies for tomorrow's world: Volume 1: Analysis. OECD.

OECD. (2016). PISA 2015 Assessment and analytical framework. Science, reading, mathematic and fi nancial literacy. OECD.

Osborne, J. (2013). The 21st century challenge for science education: Assessing scientific reasoning. *Thinking Skills and Creativity, 10*, 265-279. https://doi.org/10.1016/j.tsc.2013.07.006

Osterhaus, C., Koerber, S., & Sodian, B. (2017). Scientific thinking in elementary school: Children's social cognition and their epistemological understanding promote experimentation skills. *Developmental Psychology, 53*, 450-462. https://doi.org/10.1037/dev0000260

Osterhaus, C., Koerber, S., & Sodian, B. (2020). The Science-P Reasoning Inventory (SPR-I): Measuring emerging scientific-reasoning skills in primary school. *International Journal of Science Education, 42*, 1087-1107.

https://doi.org/10.1080/09500693.2020.1748251

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*, 543-578.

Pajares, F., & Schunk, D.H. (2001). The development of academic self-efficacy. In A. Wigfield & J.S. Eccles (Eds.), *Development of achievement motivation* (pp. 15-31). Academic Press.

Pajares, F., & Valiante, G. (1997). Influence of Self-Efficacy on Elementary Students' Writing. *The Journal of Educational Research, 90*, 353-360. https://doi.org/10.1080/00220671.1997.10544593

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology, 34*, 29-48. https://doi.org/10.1080/01443410.2013.797339

Pintrich, P. R., Smith, D. A. F., García, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. National Center for Research to Improve Postsecondary Teaching and Learning.

Pollmeier, J., Hardy, I., Koerber, S., & Möller, K. (2011). Lassen sich naturwissenschaftliche Lernstände im Grundschulalter mit schriftlichen Aufgaben valide erfassen? *Zeitschrift Für Pädagogik, 57*, 834-853.

Pollmeier, J., Tröbst, S., Hardy, I., Möller, K., Kleickmann, T., Jurecka, A., & Schwippert, K. (2017). Science-P I: Modeling conceptual understanding in primary school. In D. Leutner, J. Fleischer, J. Grünkorn, & E. Klieme (Eds.), *Methodology of Educational Measurement and Assessment. Competence Assessment in Education* (pp. 9-17). Springer International Publishing.

Ramnarain, U., & Ramaila, S. (2018). The relationship between chemistry self-efficacy of South African first year university students and their academic performance. *Chemistry Education Research and Practice*, *19*, 60-67. https://doi.org/10.1039/C7RP00110J

Ramseger, J. (2009). Experimente, Experimente. Was lernen Kinder im naturwissenschaftlichen Unterricht. *Die Grundschulzeitschrift, 225*, 14–20.

Retrieved from http://www.bildungsplaene-bw.de/site/bildungsplan/get/documents/lsbw/export-pdf/depot-pdf/ALLG/BP2016BW_ALLG_GS_SU.pdf

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*, 353-387. https://doi.org/10.1037/a0026838

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*, 1-36. https://doi.org/ 10.18637/jss.v048.i02

Schunk, D. H. (1995). Self-efficacy and education and instruction. In J. E. Maddux (Ed.), *Self efficacy, adaptation, and adjustment: Theory, research, and application* (pp. 281-303). Plenum Press.

Schwarzer, R., & Jerusalem, M. (2002). Das Konzept der Selbstwirksamkeit. *Zeitschrift Für Pädagogik, 44*, 28-53.

Schwichow, M., Osterhaus, C., & Edelsbrunner, P. A. The relation between the control-of-variables strategy and content knowledge in physics in secondary school. *Contemporary Educational Psychology, 63*, 101923. https://doi.org/10.1016/j.cedpsych.2020.101923

Siefer, K., Leuders, T., & Obersteiner, A. (2020). Leistung und Selbstwirksamkeitserwartung als Kompetenzdimensionen: Eine Erfassung individueller Ausprägungen im Themenbereich lineare Funktionen. *Journal Für Mathematik-Didaktik, 41*, 267-299. https://doi.org/10.1007/s13138-019-00147-x

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology, 99*, 611-625. https://doi.org/10.1037/0022-0663.99.3.611

Sodian, B. (2018). The development of scientific thinking in preschool and elementary school age: A conceptual model. In F. Fischer, C. A. Chinn, K. Engelmann, & J. Osborne (Eds.), *Scientific reasoning and argumentation: The roles of domain-specific and domain-general knowledge* (pp. 227-250). Routledge.

Sodian, B., & Koerber, S. (2007). Entwicklung des natunruissenschaftlichen Denkens. In Kahlert, J., Fölling-Albers, M, M. Götz, Hartinger A, D. von Reeken, & Wittkowske S. (Eds.), *Handbuch Didaktik des Sachunterrichts* (pp. 348-353). Klinkhardt.

Songer, N. B., & Linn, M. C. (1991). How do students' views of science influence knowledge integration?. *Journal of Research in Science Teaching, 28*, 761-784. https://doi.org/10.1002/tea.3660280905

Stathopoulou, C., & Vosniadou, S. (2007). Exploring the relationship between physics-related epistemological beliefs and physics understanding. *Contemporary Educational Psychology, 32*, 255-281. https://doi.org/10.1016/j.cedpsych.2005.12.002

Stender A., Schwichow, M., Zimmerman, C., & Härtig, H. (2018). Making inquiry-based science learning visible: The influence of CVS and cognitive skills on content knowledge learning in guided inquiry. *International Journal of Science Education, 40*, 1812-1831. https://doi.org/10.1080/09500693.2018.1504346

Stipek, D. C., & Hoffman, J. H. (1980). Children's achievement-related expectancies as a function of academic performance histories and sex. *Journal of Educational Psychology, 70*, 154-166.

Usher, E. L., & Pajares, F. (2009). Sources of self-efficacy in mathematics: a validation study. *Contemporary Educational Psychology, 34*, 89-101. https://doi.org/10.1016/j.cedpsych.2008.09.002

van der Graaf, J., Segers, E., & Verhoeven, L. (2015). Scientific reasoning abilities in kindergarten: Dynamic assessment of the control of variables strategy. *Instructional Science, 43*, 381-400. https://doi.org/10.1007/s11251-015-9344-y

Zumbrunn, S., Broda, M., Varier, D., & Conklin, S. (2020). Examining the multidimensional role of self-efficacy for writing on student writing self-regulation and grades in elementary and high school. *British Journal of Educational Psychology, 90*, 580-603. https://doi.org/10.1111/bjep.12315