Metacognition and Learning (2022) 17:837–854
https://doi.org/10.1007/s11409-022-09324-9

Examining the development of metacognitive strategy knowledge and its link to strategy application in complex problem solving – a longitudinal analysis

Björn Nicolay[1] · Florian Krieger[2] · Matthias Stadler[3] · Mari-Pauliina Vainikainen[4] · Marlit Annalena Lindner[5] · Anne Hansen[6] · Samuel Greiff[1]

Received: 2 July 2021 / Accepted: 27 September 2022 / Published online: 14 October 2022

Abstract
The ability to solve complex problems successfully represents a key competence for students' educational success and beyond. While strategy application and metastrategic knowledge constitute two underlying components that drive successful complex problem solving (CPS), little is known about how these two facets develop individually and jointly in students over time. In order to address this critical research gap, the present study employed a longitudinal design investigating how strategy application, with a focus on the vary-one-thing-at-a-time (VOTAT) strategy, and metastrategic knowledge evolve in students from grade 6 (t1; age $M = 12.22$) to grade 9 (t2; age $M = 15.27$). At both measurement occasions, $N = 918$ students completed two computer-based assessments, one for CPS VOTAT application, and the other for metastrategic knowledge, each consisting of six items. While initial analyses yielded statistically significant improvements in VOTAT application and metastrategic knowledge from t1 to t2, students appeared to be far from mastering either at both measurement occasions. Furthermore, results from a cross-lagged panel model showed that the two concepts are closely intertwined and mutually influence each other over time. Implications of this mutual development of VOTAT application and metastrategic knowledge in CPS are illustrated with respect to potential applications in educational contexts. The discussion places particular emphasis on how upcoming CPS training programs in educational settings can be tailored to specifically improve both strategy application and metastrategic knowledge in students.

Keywords Complex problem solving · Strategy application · Metastrategic knowledge · Development · Education

Over the course of one's educational journey, from preschool up to the end of university studies and beyond, the ability to accurately plan, organize, evaluate, and adapt one's learning approach becomes increasingly important (e.g., Stadler et al., 2018). In today's complex and dynamic educational environment, educational institutions and stakeholders are tasked with equipping their

✉ Björn Nicolay
bjorn.nicolay@uni.lu

Extended author information available on the last page of the article

students with the skills they need to successfully navigate and adapt to these demanding and changing circumstances. In turn, given the increasing prevalence of non-routine and interactive tasks at the workplace (Autor et al., 2003; Neubert et al., 2015), the ability to overcome complex challenges also helps facilitate students' future occupational success, career advancement, and lifelong learning efforts (Mainert et al., 2015, 2019). One way in which this ability can be fostered among students are training programs targeting educationally relevant skills (Eccles & Feltovich, 2008; Zhang et al., 2017). Complex problem solving (CPS) is a pertinent example of such a skill, as it addresses the challenging, rapidly changing circumstances students face nowadays (Kim et al., 2013; Liu & Liu, 2020). As such, existing research has underlined the overlap of relevant strategies between CPS and proper scientific thinking (e.g., Kuhn et al., 2008) and the well-established benefit of CPS for success in education and beyond (Jamaludin & Hung, 2017; Lotz et al., 2016; Schoenfeld, 2017; Sonnleitner et al., 2013).

As existing CPS training programs in educational contexts have largely failed to yield long-term improvements in CPS performance and transfer tasks (Kretzschmar & Süß, 2015; see also Dörner & Funke, 2017), how CPS training programs should be designed in order to yield these desired outcomes remains an open question. On the one hand, previous research has shown that strategy application plays a key role in successful CPS performance (Greiff et al., 2015b; Lotz et al., 2017; Mustafić et al., 2019). On the other hand, existing studies suggest that, in addition to strategy application, metacognitive aspects including metastrategic knowledge may be crucial for solving complex problems successfully (Molnár & Csapó, 2018; Stadler et al., 2019a.

According to Flavell (1976), a pioneering researcher who investigated the facets of cognitive development, metacognition can be defined as "one's knowledge concerning one's own cognitive processes and products or anything related to them" (p. 232). Many existing studies have addressed relevance of metacognition in the educational context, ranging from its particular beneficial effects for student performance in Science, Technology, Engineering, and Mathematics (STEM) subjects to its perks for overarching educational success (Azevedo & Aleven, 2013; Chatzipanteli et al., 2014; Dori et al., 2018; Ohtani & Hisasaka, 2018; Vestal et al., 2017; Wang et al., 2021; Zohar & Barzilai, 2013).

Previous research has indicated that CPS competency evolves gradually over time in adolescents (Frischkorn et al., 2014). In addition, firm links between students' systematic strategy application as well as metacognitive skills and overall CPS performance have already been established (Greiff et al., 2015b; Rudolph et al., 2017). However, at present, our knowledge about how these two specific cornerstones of successful CPS evolve individually and potentially interact over time is scarce at best. Therefore, the present study investigates the development of metastrategic knowledge and related strategy application in the context of CPS in a longitudinal design in order to inform assessment and training of both concepts, with the ultimate aim of successfully fostering students' overarching CPS skills in the educational context.

Strategy application in complex problem solving

CPS skills refer to the ability to solve problems with particular characteristics, such as a high number of interrelated variables, hidden variable connections, dynamic and autonomous variable changes, and multiple goals to be reached (Schoppek & Fischer, 2015; Stadler et al., 2019b). Ideally, CPS approaches follow a systematic pattern of identifying and evaluating the mechanisms of the respective variables present in a given problem space one after another, in order to deduct whether and how a particular variable exerts influence on other variables (e.g., Van der Graaf et al., 2015). This strategy of manipulating one variable in isolation in order to investigate its potential effect(s) on the remaining variables is termed vary-one-thing-at-a-time (VOTAT; e.g., Greiff et al., 2015b; or the control-of-variables strategy; CVS; e.g., Schwichow et al., 2016), and its application has generally been associated with a higher chance of successful CPS performance (Molnár & Csapó, 2018; Mustafić et al., 2019; Wüstenberg et al., 2014). Alongside VOTAT's key role for CPS success, systematic application of this strategy has also been found to be relevant in the broader educational realm; for instance, as a facilitator of content learning in scientific inquiry (Chen & Klahr, 1999; Hovardas et al., 2017; Schwichow et al., 2016; Stender et al., 2018; Teig et al., 2020).

Due to the immense benefits associated with VOTAT strategy usage for both CPS and overall educational success, it is of great importance to understand how VOTAT application and awareness evolve in students. Existing studies have shown that, in CPS assessment contexts, a large proportion of students do not or only insufficiently apply VOTAT, even over the course of multiple complex problem items (Greiff et al., 2018; Lotz et al., 2017; Molnár & Csapó, 2018; Mustafić et al., 2019; Wu & Molnár, 2021; Wüstenberg et al., 2014). Thus, students apparently do not apply VOTAT inherently and intuitively, which calls for a thorough investigation of how students' VOTAT application develops over time, in order to be able to adapt future CPS training programs accordingly.

Metacognitive strategy knowledge in complex problem solving

In addition, while the mere and not necessarily conscious application of VOTAT yields a higher chance of successful CPS performance compared to an unsystematic approach (e.g., Lotz et al., 2017), it is by no means a guarantor for CPS success (Kuhn & Dean, 2005). For instance, many students apply the VOTAT strategy as intended, yet still fail to solve a given complex problem successfully (Stadler et al., 2019a). Moreover, in their empirical study, Molnár and Csapó (2018) discovered that students who intentionally applied VOTAT (i.e., 'conscious VOTAT strategy users') outperformed their counterparts who also used VOTAT, but did so unintentionally (i.e., 'non-conscious VOTAT strategy users'). This finding suggests that not only the ability to apply VOTAT, but also metastrategic knowledge about VOTAT facilitates CPS performance. Thus, knowing how and when to use VOTAT and applying it accordingly appears to be a better precursor of CPS success than mere VOTAT application alone.

Generally, metacognitive knowledge represents one key facet of metacognition (Flavell, 1979). Metacognitive knowledge can be defined as explicit knowledge about one's cognitive processes and the results thereof (Efklides, 2011), and includes the sub-facet strategic knowledge, also termed metastrategic knowledge, which, in turn, refers to declarative knowledge about the particular benefits and downsides of applying a particular strategy in a given situation (Jia et al., 2019). Thus, in the context of CPS, the presence of metastrategic knowledge in students might be manifested by them applying VOTAT in a conscious and deliberate fashion.

Previous research has unanimously indicated the benefits of metastrategic knowledge for CPS success (Molnár & Csapó, 2018; Wüstenberg et al., 2014). In this regard, a study by Wüstenberg et al. (2014) serves as a direct precursor to the present research. In their study, the authors predicted students' VOTAT application from fluid intelligence, scientific reasoning, and learning orientation, resulting in a considerable amount of variance in VOTAT application remaining unexplained. Based on their findings, Wüstenberg et al. (2014) conclude that metastrategic knowledge requires further scientific scrutiny as a promising candidate to play a key role in VOTAT application. Yet, currently, there is still a dearth of research on how and whether metastrategic knowledge and strategy application interact as well as evolve individually and/or in conjunction with one another over time in students. Extending our knowledge with regard to the independent and potentially reciprocal developmental mechanisms behind both these aspects as precursors of successful CPS is crucial in order to optimize upcoming CPS training programs in educational contexts, particularly in light of the fact that existing CPS training programs generally have reported insufficient training and transfer effects (Kretzschmar & Süß, 2015).

The present study

Given the current state of available research and research gaps with regard to the joint role of strategy application and metastrategic knowledge for successful CPS performance, the aim of the present study is to investigate how both strategy (i.e., VOTAT) application and metastrategic knowledge develop, influence each other, and change over time in students. While previous research efforts have paved the way for the present study by showing that VOTAT application does not automatically constitute successful CPS performance, they exclusively rely on cross-sectional data (Molnár & Csapó, 2018; Stadler et al., 2019a). In contrast, the present study employs a longitudinal design incorporating two measurement occasions, which represents a particular added value, as it allows for capturing the development of both strategy application and metastrategic knowledge over time.

Furthermore, by analyzing the temporal evolution of both these relevant facets for CPS performance simultaneously, we address additional pertinent research gaps, including the underrepresentation of domain-general constructs, including problem solving, in metacognition research, as well as the comparative lack of studies on the development of metastrategic knowledge in students (Zohar & Barzilai, 2013). Importantly, as suggested by previous studies in related fields, metastrategic knowledge represents an important skill for CPS training programs in educational contexts aiming at eliciting long-term and transfer effects to address (Jia et al., 2019; Montague, 1991; Zumbach et al., 2020), particularly given that this type of knowledge generally does not evolve automatically in students (Karlen et al., 2014).

As we are exploring uncharted territory with regard to the developmental interplay of strategy application and metastrategic knowledge in CPS, different possible scenarios for their potential mutual influences exist. Firstly, strategy application and metastrategic knowledge might develop independently over time. Secondly, initial VOTAT application may influence the subsequent development of metastrategic knowledge about VOTAT. Thirdly, students may accumulate metastrategic knowledge about VOTAT prior to being able to apply VOTAT. Taking into account these multiple potential scenarios, the present study aims at addressing the following three research questions (RQs):

1. How does strategy (i.e., VOTAT) application develop in students over time?
2. How does metastrategic knowledge about VOTAT develop in students over time?
3. How do the developmental trajectories of VOTAT application and metastrategic knowledge about VOTAT interact over time in students?

Materials and methods

Sample characteristics

The sample consisted of $N = 1{,}316$ grade six students ($n = 646$ females, $n = 35$ missing information) at the first data collection wave (t1), with a mean age of 12.22 years ($SD = 0.44$). The second data collection wave (t2) took place three years later, when the same students were in grade nine (age $M = 15.27$, $SD = 0.46$). After filtering (please see Section "Filtering and Preliminary Analysis" below for further details), the final sample size was $N = 918$. All data were collected from students in a Finnish municipality. Importantly, the sample's representativeness concerning several characteristics, including demographics, is ensured (Vainikainen, 2014). All students needed to provide informed consent by agreeing to answer the questions and complete the tasks before they could begin participation.

Materials

Assessment of VOTAT application

VOTAT application was assessed over the course of five CPS items from the MicroDYN assessment framework (Greiff et al., 2012). The MicroDYN approach employs a multitude of different CPS items with arbitrary cover stories and similar underlying features based on the principle of linear structural equations (Funke, 2001). Please see Fig. 1 below for a sample item.

MicroDYN incorporates the two CPS phases of knowledge acquisition and knowledge application (e.g., Greiff et al., 2015a). Within a given problem space, several input (see left part of Fig. 1) and output variables (see right part of Fig. 1) are presented. In the initial knowledge acquisition phase, the user is asked to manipulate the input variables in order to detect their initially hidden effects on the output variables. This can be achieved by increasing or decreasing the dosages of the input variables and subsequently clicking on 'Apply', which potentially triggers a change in the values of the output variables. Whenever a relationship between an input variable and output variable(s) has been discovered, the solver is required to plot said relationship in the visual model of the problem space (see bottom part of Fig. 1), by clicking on both the input and output variable in order to draw a blue arrow between the two indicating the presence of a relationship. Subsequently, in the knowledge application phase, all output variables contain an initial value and a target value area, which should be reached in as few steps as possible by manipulating the input variables accordingly.

Assessment of metastrategic knowledge about VOTAT

Metastrategic knowledge about VOTAT was assessed using an updated version (Hautamäki et al., 2002) of the 'Pendulum' task originally created by Shayer (1976) to measure students' formal operational thinking skills using the control-of-variables (i.e., VOTAT) strategy. Students were asked to complete a total of six items involving comparison sets containing four input variables ('driver', 'car', 'tires', and 'track') and one output variable (lap time) in a fictitious Formula 1 environment. Over the course of the first four tasks, students were presented with different scenarios in which only one or multiple input variables were varied and the rest held constant. In each item, students were asked to judge whether it was possible to infer the

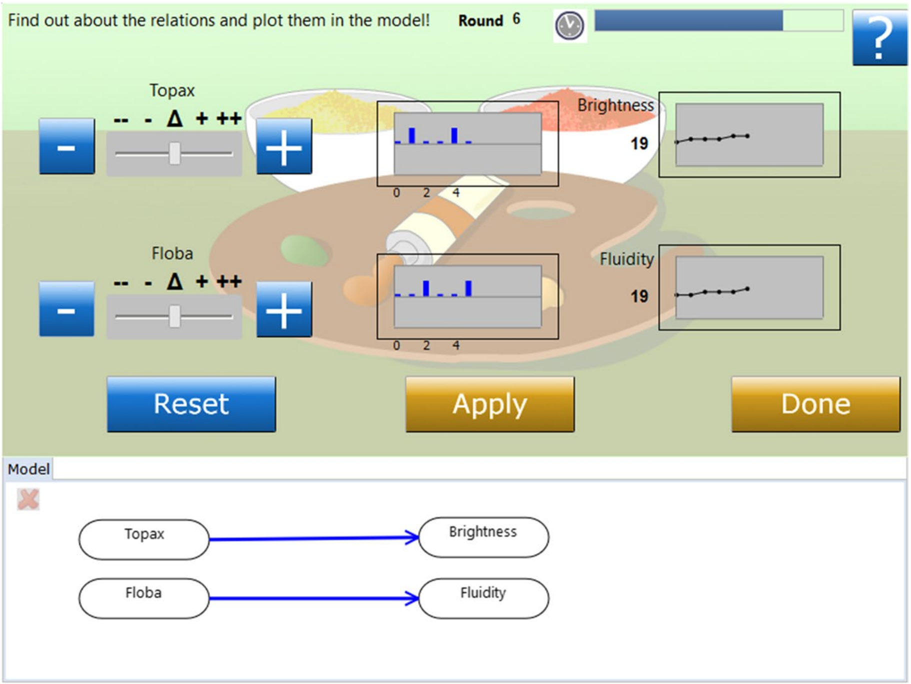

Fig. 1 Problem space (top) featuring two input variables (left; e.g., 'Topax') and two output variables (right; e.g., 'Fluidity'), and visual model of variable relationships (bottom) from the MicroDYN item "Drawing"

impact of a particular input variable on lap time with certainty and accuracy based on the given scenario. The remaining two tasks required the students to indicate which input variables to vary in order to reach a predefined goal, such as evaluating the particular effect of the input variable 'car' on the output variable lap time. For each item, students' responses were scored dichotomously (0 = failure; 1 = success). Figure 2 below depicts a sample item from the assessment of metastrategic knowledge about VOTAT.

Procedure

Both the CPS and the metastrategic knowledge assessment were part of a larger computer-based test battery including additional assessment instruments for reading comprehension and working memory, amongst other variables, and were completed on computers in class. Students first completed the metastrategic knowledge items as part of this larger overarching test battery, which took about 90 min in total. The CPS assessment took place one week later, and students completed nine MicroDYN items in a predefined order (total test time 45 min).

Scoring

Due to the computer-based administration of the CPS tasks, all actions (i.e., input variable manipulations) taken by each participant during the MicroDYN items were stored

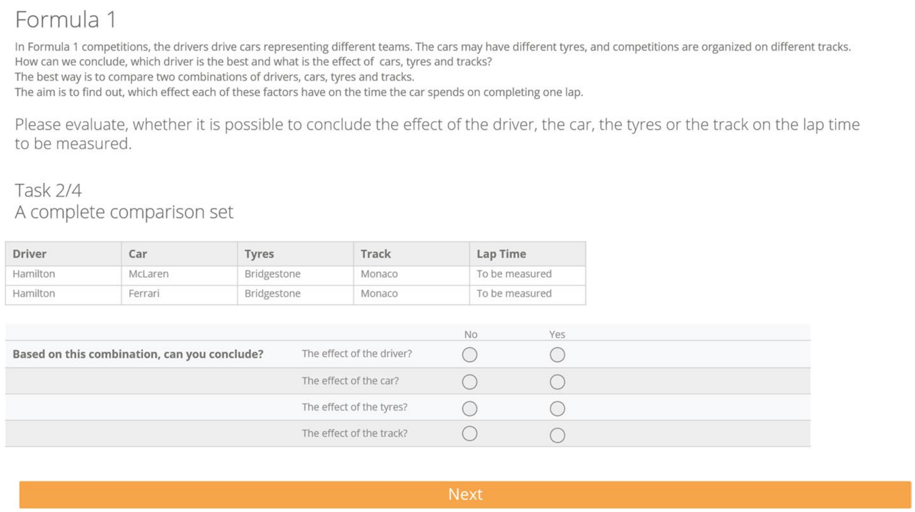

Fig. 2 Sample task from metastrategic knowledge assessment featuring a fictitious Formula 1 environment comparison set

automatically in XML log files. For each MicroDYN item, students' VOTAT application was scored dichotomously (0 = no VOTAT application; 1 = VOTAT was applied at least once). Based on the item-based dichotomous scores for each individual, we created a VOTAT application sum score across all items for the subsequent paired-sample t-test analysis. For our cross-lagged panel analysis (see Sect. "RQ3: How do the developmental trajectories of VOTAT application and metacognitive strategy knowledge about VOTAT interact over time?" below), however, we retained each dichotomous item-based score as a manifest variable loading onto a latent overarching VOTAT application variable.

With regard to the items assessing students' metastrategic knowledge, students received a passing score if they judged a given scenario correctly and a failing score if they misjudged a given scenario. Based on their individual item scores, a cumulative score was calculated for each student, with values ranging from 0 (no items solved correctly) to 6 (all items solved correctly). This score was then used as a single manifest variable for further statistical analysis.

Statistical analysis

Filtering and preliminary analysis

Overall, the CPS test consisted of nine consecutive MicroDYN items, six of which we retained for statistical analyses (i.e., 'Lemonade', ''Drawing', 'Cat', 'Moped', 'Game', and 'Handball'). The remaining three MicroDYN items (i.e., 'Gardening', 'Spaceship', and 'First Aid') were removed because they possessed an additional feature known as eigendynamic, which requires the solver to apply another strategy in addition to VOTAT in order to solve the item successfully, thereby tainting a clear-cut investigation of VOTAT application (Lotz et al., 2017; Schoppek & Fischer, 2017; Stadler et al., 2016). Upon closer inspection, we discovered that VOTAT was not applied at all in any of the first three MicroDYN items by students in grade 6. Thus, these three items (i.e., 'Lemonade', ''Drawing', and 'Cat') were removed from the cross-lagged panel model

(CLPM) analysis for these students, based on their missing variance. However, the items were retained for the VOTAT application sum score in the respective paired samples t-tests as well as for all further analyses involving the students in grade 9, as these students provided a variance of VOTAT application throughout all six MicroDYN items. Furthermore, of the $N = 1,316$ students, some did not complete all six MicroDYN items, while others did not provide data at both time points, resulting in several missing values ($n = 398$). In order to analyze only students with full valid response sets on both occasions, the ones with missing values were not considered in the further analyses. After removing all said students, the final sample size used for all subsequent analyses was $N = 918$.

Version 0.14.0.0 of the statistical software JASP (JASP Team, 2020) was used for the initial demographical analyses as well as to evaluate the development of both VOTAT application (RQ1) and metastrategic knowledge about VOTAT (RQ2) individually over time by means of two separate paired samples t-tests (see Sects. 3.1 and 3.2, respectively).

First, we investigated several descriptive statistics for our four key variables VOTAT application in grades six (t1) and nine (t2), respectively, and metastrategic knowledge about VOTAT in grades six (t1) and nine (t2), respectively. The results are summarized in Table 1 below.

Cross-Lagged Panel Model (CLPM)

In order to assess the joint developmental trajectories of students' VOTAT application and metastrategic knowledge over time (RQ3), we used version 8.6 of the statistical software MPlus (Muthén & Muthén, 2021). We specified a CLPM (Kenny, 1975, 2014), incorporating the cross-lagged prediction paths between VOTAT application and metastrategic knowledge about VOTAT over time, the linear prediction paths of VOTAT application and metastrategic knowledge about VOTAT over time, and finally, the correlations between VOTAT application and metastrategic knowledge about VOTAT at both t1 (grade 6) and t2 (grade 9).

Results

RQ1: How does VOTAT application develop over time?

Firstly, we investigated whether to what extent VOTAT application rates changed in students between t1 and t2 (i.e., from 6 to 9th grade) using the VOTAT application sum score. Results

Table 1 Means, standard deviations, and correlations with confidence intervals for key variables

Variable	*M*	*SD*	1	2	3
1. VOTAT Application Grade 6 (t1)	1.06	1.32			
2. VOTAT Application Grade 9 (t2)	2.89	2.56	0.42* [0.37, 0.48]		
3. Metatstrategic Knowledge Grade 6 (t1)	2.18	1.60	0.41* [0.36, 0.47]	0.35* [0.29, 0.41]	
4. Metatstrategic Knowledge Grade 9 (t2)	2.68	1.57	0.35* [0.29, 0.41]	0.45* [0.40, 0.50]	0.47* [0.41, 0.52]

$N = 918$. *M* and *SD* represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. * indicates $p < 0.01$

of a paired samples t-test showed that students' VOTAT application rates were statistically significantly higher in 9th ($M=2.89$, $SD=2.56$) compared to 6th grade ($M=1.06$, $SD=1.32$), with $t(917)=23.83$ ($p<0.001$); $d=0.79$.

RQ2: How does metastrategic knowledge about VOTAT develop over time?

Secondly, we carried out another paired samples t-test to evaluate the development of metastrategic knowledge about VOTAT in students over time using the metastrategic knowledge sum score. The results indicated that students' metastrategic knowledge levels were statistically significantly lower in grade six ($M=2.18$, $SD=1.60$) compared to grade nine ($M=2.68$, $SD=1.57$), with $t(917)=9.19$ ($p<0.001$); $d=0.30$.

RQ3: How do the developmental trajectories of VOTAT application and metacognitive strategy knowledge about VOTAT interact over time?

In order to evaluate the joint development of VOTAT application and metacognitive strategy knowledge, we specified the CLPM as described in Section "Cross-Lagged Panel Model" above. The results showed that our CLPM fit the data well (according to conventions proposed by Hu & Bentler, 1999), with $\chi^2=131.229$; $df=40$; $p<0.001$; RMSEA = 0.050; SRMR = 0.021; CFI = 0.997; TLI = 0.996. As anticipated, the longitudinal paths for VOTAT application and metastrategic knowledge were statistically significant in the sense that both VOTAT application ($\beta=0.494$) and metastrategic knowledge ($\beta=0.336$) at t1 predicted that same variable at t2 (both $p<0.001$). In addition, both cross-lagged paths were statistically significant. More specifically, VOTAT application at t1 predicted metastrategic knowledge at t2 ($\beta=0.266$, $p<0.001$), and metastrategic knowledge at t1 was a significant precursor of VOTAT application at t2 ($\beta=0.150$, $p<0.001$). Furthermore, VOTAT application and metastrategic knowledge were statistically significantly related to each other at both time points, with $\beta=0.489$, $p<0.001$ at t1, and $\beta=0.315$, $p<0.001$ at t2. All results of our CLPM pertaining to the interaction between the developmental trajectories of VOTAT application and metacognitive strategy knowledge about VOTAT in students over time are visualized in Fig. 3 below.

Discussion

The development of VOTAT application in students over time

First of all, VOTAT application rates increased in students from grade 6 to grade 9. Thus, as students grow older, they become more proficient in using the VOTAT strategy. This observation is in line with previous research showing that CPS skills, of which VOTAT application represents an integral part, develop gradually over time in students (e.g., Frischkorn et al., 2014). In this regard, the time span from grade 6 to 9 represents a particularly sensitive period in the advancement of students' CPS skills (Molnár et al., 2013). Our results also highlighted the importance of this time span for the development of CPS competence (i.e., VOTAT application) in our sample.

However, it also needs to be noted that, on average, students applied VOTAT as only to a very limited extend, both in grade 6 and grade 9. Due to the similar underlying features of the individual items, the MicroDYN assessment approach heavily emphasizes VOTAT application, and students are hence expected to apply VOTAT multiple times in each item. However, our results indicate that most students only applied VOTAT in about one out of six (grade 6) and in about three out of six items (grade 9), respectively. Thus, while the older students generally came closer to employing the most promising way of solving complex problems, the majority were still far from consistently applying VOTAT in all items, leaving considerable room for improvement in this regard. As such, 58% of students in grade 6 refrained from applying VOTAT in general (with an additional 15% applying VOTAT in up to two items). In contrast, only 34% of students in grade 9 did not apply VOTAT at all (with 14% of students applying VOTAT in up to two items). Likewise, notable percentages of students either did not apply VOTAT at all or only to a very limited extent in previous studies (e.g., Molnár & Csapó, 2018).

The development of metastrategic knowledge in students over time

Secondly, and akin to VOTAT application, students' metastrategic knowledge also increased from grade 6 to grade 9. Again, these findings can easily be integrated into the current picture of how metastrategic knowledge evolves in students. As discussed in previous research, the development of metastrategic knowledge usually progresses gradually in students from elementary up to high school (Veenman et al., 2004; Zimmerman, 2007). Interestingly, some studies even argue for the presence of metastrategic knowledge in students as early as second grade (e.g., Luwel et al, 2003).

While we were able to confirm such developmental progress in students with regard to metastrategic knowledge, our results also showed that, on average, students were only able to solve about one third (grade 6) to about half of the items (grade 9) related to metastrategic knowledge correctly. Thus, there appears to be additional room for further and more elaborate development of metastrategic knowledge beyond the levels observed in the students attending grade 9. However, as mentioned before, existing research suggests that metastrategic knowledge evolves in students over a considerable period of time from elementary up to high school (Veenman et al., 2004; Zimmerman, 2007), which means that we would not expect it to be already or even nearly "fully" developed in our sample of students at t2. This notion is further underlined when considering that, in the present study,

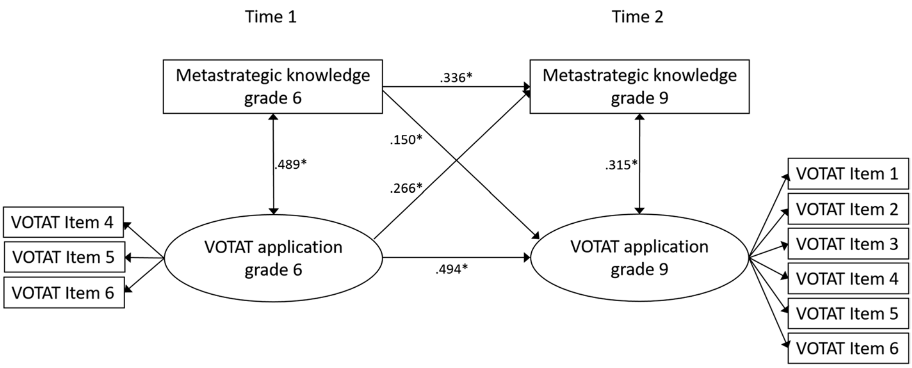

Fig. 3 Results of cross-lagged panel model investigating the joint development and individual predictive reciprocal relationships between VOTAT application and metastrategic knowledge in students from Grade 6 (Time 1) to Grade 9 (Time 2). *Note.* $N = 918$. Values represent standardized coefficients. * indicates that $p < 0.001$

the metastrategic knowledge tasks were part of a larger test battery assessing formal operational thinking. Although the early years of secondary school mark a crucial period for students' cognitive development, only a few students have reached this stage by age 12, and according to the literature, a considerable proportion of them do not seem to have mastered formal operational thinking even at age 15 (Ginsburg & Opper, 1988; Inhelder & Piaget, 1958; Khoirina & Cari, 2018; Lawson, 1978).

Overall, when evaluated individually, our results reveal similar developmental trajectories for both VOTAT application and metastrategic knowledge thereof from grade 6 to grade 9 – students show improvement in both, albeit leaving sufficient room for further improvement before achieving mastery of either skill. Importantly, these findings highlight the importance of facilitating both CPS facets through CPS training programs, so that students become proficient in applying VOTAT and employing their metastrategic knowledge sooner.

The developmental interplay of VOTAT application and metastrategic knowledge in students

After investigating how VOTAT application and metastrategic knowledge evolve individually in students, we evaluated the developmental interplay of both CPS facets over time by means of a CLPM. Overall, our results point toward a reciprocal relationship between VOTAT application and metastrategic knowledge (see Fig. 3). Hence, the two concepts appear to be strongly intertwined, mutually influence each other, and allegedly develop in tandem over time (Kuhn & Pearsall, 1998).

From a practical point of view, this simultaneous and reciprocal development of both VOTAT application and metastrategic knowledge seems straightforward. Students are continuously exposed to the VOTAT strategy during their regular school instruction (e.g., Schwichow et al., 2016), and eventually learn to apply this strategy with ease. In addition, when prompted to monitor or evaluate their way of approaching and working on a given problem (e.g., solving a mathematical equation), students might actively engage in consolidating and expanding their metastrategic knowledge base. This procedure could also be applied in reverse, for example, by having students learn about VOTAT application theoretically first before actively practicing it. Generally, the facilitation of explicit links between theory and practice as well as their reciprocal consolidation represents a hallmark of contemporary education (e.g., Shaharabani & Yarden, 2019).

Furthermore, we acknowledge that some 'natural' development of VOTAT application and metastrategic knowledge over time is present in our data, even without students' receiving explicit CPS training. This observation illustrates that both VOTAT application and metastrategic knowledge about VOTAT may not be limited to CPS, despite being inherently linked to it, but rather represent domain-general skills that are learned and applied continuously by students across multiple subjects, particularly in the STEM domain (Dori et al., 2018; Schwichow et al., 2016; Wang et al., 2021). As a result, we do not posit that a CPS training program is strictly necessary for students to learn the basic principles of scientific thinking and eventually reach the formal operational stage (e.g., Inhelder & Piaget, 1958). However, previous research indicates that the seemingly natural development of VOTAT application and metastrategic knowledge about VOTAT as underlying facets of proper scientific thinking can be significantly facilitated and extended by a comprehensive CPS training program (e.g., Adey et al., 2007). Hence, and taking into account the generally developable nature of students' proficiency in VOTAT application and metastrategic knowledge about VOTAT, a CPS training program represents a promising method for further systematically enhancing students' proficiency in this area.

Main implications and limitations

Our results have several noteworthy implications for the trainability of both VOTAT application and students' metastrategic knowledge with respect to CPS in educational contexts. Firstly, while students' VOTAT application and metastrategic knowledge evolve to a certain extent from grade 6 to grade 9 without deliberate CPS training, both components still appear to be far from fully developed in grade 9. Thus, to actively foster VOTAT application and metastrategic knowledge (i.e., CPS competence) beyond what the ordinary curriculum offers, students may particularly benefit from a tailored CPS training program.

Secondly, upcoming CPS training programs should ideally seek to deliberately train both VOTAT application as well as students' metastrategic knowledge. Therefore, CPS training programs that merely tell students to work on given CPS tasks and apply VOTAT may fall short of facilitating students' metastrategic knowledge and thereby produce limited transfer effects (see Kretzschmar & Süß, 2015). Furthermore, as stated by Kuhn and Pearsall (1998), strategy application appears to be difficult to master without students possessing a profound level of metastrategic knowledge. Previous training programs targeting both components in fields adjacent to CPS have yielded notable improvements in both strategy application and metastrategic knowledge among students (e.g., Kuhn & Pearsall, 1998).

While the question how to train VOTAT application can be answered in a straightforward fashion (i.e., telling students to apply this strategy over the course of multiple CPS items), we would like to discuss the possibilities for fostering students' metastrategic knowledge in CPS training programs in more detail. To begin with, prior research has unanimously shown that training programs for metastrategic knowledge can be beneficial, particularly for initially low-performing students (Zohar, 2012). This beneficial effect has been found not only in laboratory experimental designs, but also in classroom intervention training programs (Zohar & Ben David, 2008). As argued in the systematic review by Schwichow et al. (2016), one promising approach to facilitate students' metastrategic knowledge, especially with regard to VOTAT, is inducing cognitive conflict. Cognitive conflict targets the modal validity of students' approaches to working on a particular problem and requires them to evaluate whether what they are doing at a given moment constitutes a valid and unconfounded scientific approach (Schwichow et al., 2016). In transferring this notion to practice, teachers or researchers could ask students at a given point during their work on a CPS item if they will be able to draw scientifically sound inferences from their procedure (i.e., using VOTAT), irrespective of context-specific item characteristics such as variable names. Hence, a CPS training emphasizing such an overarching monitoring process represents a promising way to promote students' metastrategic knowledge along with their overall CPS competence, which has received some preliminary empirical support from existing research (e.g., An & Cao, 2014). Moreover, the act of predicting a given outcome has been termed a unique learning strategy in recent educational research by Brod (2021). Interestingly, the results of Brod's (2021) study point to the particular usefulness of incorrect predictions, which induce higher attention to the correct solution and how the initial misconception came about. Thus, upcoming CPS training programs might incorporate this aspect, for instance by showing students different video segments of someone working on a complex problem and asking them about the anticipated outcome (e.g., will the problem be solved correctly or incorrectly and why?). Subsequently, a reflection could take place in which students evaluate whether and why their prediction was correct or incorrect. By including these aspects, CPS training programs may further foster students' awareness of how systematic strategy application (particularly VOTAT) leads to successful CPS performance. Simultaneously, making incorrect predictions may help counteract students' possible overconfidence in their CPS abilities, which has been shown to result in weaker performance in computer-based science simulations comparable to contemporary CPS assessment approaches (e.g., Finn, 2018).

At this point, we would like to address some limitations of the present study, as well as some suggestions for future research. Firstly, the application of CLPM to derive meaningful scientific inferences has been criticized in recent years, and some alternative advanced procedures have been presented (Hamaker et al., 2015; Mund & Nestler, 2019). However, since we only have data from two different measurement occasions at our disposal, and as the CLPM is able to account for both autoregressive and cross-lagged paths, which we were particularly interested in, the CLPM remains the preferable means of analysis in the present study (Mund & Nestler, 2019). However, future studies, ideally relying on three or more measurement occasions and thereby increasing the causal interpretability of both VOTAT application and metastrategic knowledge development in students, should apply alternative statistical approaches to the 'classical' CLPM employed here. Secondly, as argued by Zohar (2012), in addition to knowledge about strategies, metastrategic knowledge also incorporates knowledge about tasks (i.e., discerning individual characteristics of a given item that require the application of a particular strategy at a specific point in time). However, due to its underlying focus on VOTAT as a key strategy for CPS success, the MicroDYN framework does not incorporate great variance in task characteristics across items. Thus, for a thorough overarching assessment of metastrategic knowledge in CPS, upcoming studies should consider alternative CPS assessment approaches with a broader range of beneficial strategies. However, as the VOTAT strategy is crucial for unconfounded scientific experimentation and interpretation, in our opinion, it represents the logical strategy of choice to be investigated in order to foster students' CPS and scientific inquiry skills (Chen & Klahr, 1999; Greiff et al., 2015b, 2018; Stender et al., 2018; Teig et al., 2020). Thirdly, the present study did not account for the conceptual distinction between VOTAT usage per task vs. per input variable (see, e.g., Wu & Molnár, 2021). While coding VOTAT per input variable would clearly allow a more fine-grained investigation of students' VOTAT application behavior, the primary goal of the present study was to provide a first comprehensive overview of the longitudinal developmental trajectory of VOTAT application and its counterpart metastrategic knowledge in students. Therefore, we advocate for the assessment of VOTAT application per input variable in upcoming studies evaluating the relationship between strategy application and metacognition in CPS.

Conclusion

The present study shed light on the individual and joint evolution of strategy (i.e., VOTAT) application and metastrategic knowledge in students from grade 6 to grade 9. Our results yielded a simultaneous and reciprocal developmental pattern of strategy application and metastrategic knowledge in the context of CPS. Despite significant improvement in both competencies over time, our analyses indicate that students in both grade 9 and earlier would likely benefit considerably from a tailored CPS training program that specifically incorporates both VOTAT application and the enhancement of metastrategic knowledge. On a broader level, due to the pioneering role of the present study in investigating the developmental trajectories of these two crucial CPS components over time, we were able to derive several noteworthy implications and specific implementation suggestions for upcoming CPS training programs aimed at enhancing students' skills in educational settings. Overall, the present study contributes to and advances our understanding of how two central underlying driving facets of successful CPS performance develop longitudinally in students, and how they can be successfully facilitated by suitable interventions in practice.

Acknowledgements This research was supported by a grant from the Fonds National de la Recherche Luxembourg (Luxembourgish National Research Found) [CORE 'TRIOPS'].

Funding This research was supported by a grant from the Fonds National de la Recherche Luxembourg (Luxembourgish National Research Found) [CORE 'TRIOPS'].

Data availability For any inquiries related to raw data and material used in this study, please contact Mari-Pauliina Vainikainen (mari-pauliina.vainikainen@tuni.fi).

Code availability The code used for the analyses in the present study can be made available on request. Please contact either the corresponding author Björn Nicolay (bjorn.nicolay@uni.lu) or Florian Krieger (florian.krieger@tu-dortmund.de).

Declarations

Conflicts of interest/Competing interests Samuel Greiff is one of two authors of the commercially available COMPRO-test that is based on the multiple complex systems approach, which employs the same assessment principle as MicroDYN. He receives royalty fees for COMPRO.

References

Adey, P., Csapó, B., Demetriou, A., Hautamäki, J., & Shayer, M. (2007). Can we be intelligent about intelligence?: Why education needs the concept of plastic general ability. *Educational Research Review, 2*(2), 75–97. https://doi.org/10.1016/j.edurev.2007.05.001

An, Y. J., & Cao, L. (2014). Examining the effects of metacognitive scaffolding on students' design problem solving and metacognitive skills in an online environment. *Journal of Online Learning and Teaching, 10*(4), 552–568.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics, 118*(4), 1279–1333. https://doi.org/10.1162/003355303322552801

Azevedo, R., & Aleven, V. (2013). International Handbook of Metacognition and Learning Technologies. *Springer*. https://doi.org/10.1007/978-1-4419-5546-3

Brod, G. (2021). Predicting as a learning strategy. *Psychonomic Bulletin & Review*, 1-9. https://doi.org/10.3758/s13423-021-01904-1

Chatzipanteli, A., Grammatikopoulos, V., & Gregoriadis, A. (2014). Development and evaluation of metacognition in early childhood education. *Early Child Development and Care, 184*(8), 1223–1232. https://doi.org/10.1080/03004430.2013.861456

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development, 70*, 1098–1120. https://doi.org/10.1111/1467-8624.00081

Dori, Y. J., Mevarech, Z. R., & Baker, D. R. (Eds.) (2018). *Cognition, Metacognition, and Culture in STEM Education*. Springer.

Dörner, D., & Funke, J. (2017). Complex problem solving: What it is and what it is not. *Frontiers in Psychology, 8*, 1153. https://doi.org/10.3389/fpsyg.2017.01153

Eccles, D. W., & Feltovich, P. J. (2008). Implications of domain-general "psychological support skills" for transfer of skill and acquisition of expertise. *Performance Improvement Quarterly, 21*(1), 43–60. https://doi.org/10.1002/piq.20014

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist, 46*(1), 6–25. https://doi.org/10.1080/00461520.2011.538645

Finn, B. (2018). Metacognitive evaluations during science simulations: How do ratings of confidence and understanding relate to science assessment inquiry processes? *Archives of Scientific Psychology, 6*(1), 117–129. https://doi.org/10.1037/arc0000037

Flavell, J. H. (1976). Metacognitive Aspects of Problem Solving. In L. Resnick (Ed.), *The Nature of Intelligence* (pp. 231–236). Lawrence Erlbaum Associates.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Frischkorn, G. T., Greiff, S., & Wüstenberg, S. (2014). The development of complex problem solving in adolescence: A latent growth curve analysis. *Journal of Educational Psychology, 106*(4), 1007–1020. https://doi.org/10.1037/a0037114

Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking & Reasoning, 7*(1), 69–89. https://doi.org/10.1080/13546780042000046

Ginsburg, H. P., & Opper, S. (1988). *Piaget's theory of intellectual development* (3rd ed.). Prentice-Hall, Inc.

Greiff, S., Fischer, A., Stadler, M., & Wüstenberg, S. (2015a). Assessing complex problem solving skills with multiple complex systems. *Thinking & Reasoning, 21*(3), 356–382. https://doi.org/10.1080/13546783.2014.989263

Greiff, S., Molnár, G., Martin, R., Zimmermann, J., & Csapó, B. (2018). Students' exploration strategies in computer-simulated complex problem environments: A latent class approach. *Computers & Education, 126*, 248–263. https://doi.org/10.1016/j.compedu.2018.07.013

Greiff, S., Wüstenberg, S., & Avvisati, F. (2015b). Computer-generated log-file analyses as a window into students' minds?: A showcase study based on the PISA 2012 assessment of problem solving. *Computers & Education, 91*, 92–105. https://doi.org/10.1016/j.compedu.2015.10.018

Greiff, S., Wüstenberg, S., & Funke, J. (2012). Dynamic problem solving: A new assessment perspective. *Applied Psychological Measurement, 36*(3), 189–213. https://doi.org/10.1177/0146621612439620

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods, 20*(1), 102–116. https://doi.org/10.1037/a0038889

Hautamäki, J., Arinen, P., Niemivirta, M. J., Eronen, S., Hautamäki, A., Kupiainen, S., Lindblom, B., Pakaslahti, L., Rantanen, P., & Scheinin, P. (2002). *Assessing Learning- to-Learn: A framework*. Opetushallitus.

Hovardas, T., Xenofontos, N. A., & Zacharia, Z. C. (2017). Using Virtual Labs in an Inquiry Context: The Effect of a Hypothesis Formulation Tool and an Experiment Design Tool on Students' Learning. In *Optimizing STEM education with advanced ICTs and simulations* (pp. 58–83). IGI Global.

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), 1–55. https://doi.org/10.1080/10705519909540118

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking: From childhood to adolescence.* (A. Parsons & S. Milgram, Trans.). Basic Books. https://doi.org/10.1037/10034-000

Jamaludin, A., & Hung, D. (2017). Problem-solving for STEM learning: Navigating games as narrativized problem spaces for 21st century competencies. *Research and Practice in Technology Enhanced Learning, 12*(1), 1–14. https://doi.org/10.1186/s41039-016-0038-0

JASP Team. (2020). JASP (Version 0.14.1) [Computer software].

Jia, X., Li, W., & Cao, L. (2019). The role of metacognitive components in creative thinking. *Frontiers in Psychology, 10*, 2404. https://doi.org/10.3389/fpsyg.2019.02404

Karlen, Y., Merki, K. M., & Ramseier, E. (2014). The effect of individual differences in the development of metacognitive strategy knowledge. *Instructional Science, 42*(5), 777–794. https://doi.org/10.1007/s11251-014-9314-9

Kenny, D. A. (2014). Cross-lagged panel design. *Wiley StatsRef: Statistics Reference Online*. https://doi.org/10.1002/9781118445112.stat06464

Kenny, D. A. (1975). Cross-lagged panel correlation: A test for spuriousness. *Psychological Bulletin, 82*(6), 887–903. https://doi.org/10.1037/0033-2909.82.6.887

Khoirina, M., & Cari, C. (2018). Identify students' scientific reasoning ability at senior high school. In *Journal of Physics: Conference Series* (vol. 1097, no. 1, p. 12–24). IOP Publishing. https://doi.org/10.1088/1742-6596/1097/1/012024

Kim, Y. R., Park, M. S., Moore, T. J., & Varma, S. (2013). Multiple levels of metacognition and their elicitation through complex problem-solving tasks. *The Journal of Mathematical Behavior, 32*(3), 377–396. https://doi.org/10.1016/j.jmathb.2013.04.002

Kretzschmar, A., & Süß, H. M. (2015). A study on the training of complex problem solving competence. *Journal of Dynamic Decision Making, 1*, 4. https://doi.org/10.11588/jddm.2015.1.15455

Kuhn, D., & Dean, D., Jr. (2005). Is developing scientific thinking all about learning to control variables? *Psychological Science, 16*(11), 866–870. https://doi.org/10.1111/j.1467-9280.2005.01628.x

Kuhn, D., Iordanou, K., Pease, M., & Wirkala, C. (2008). Beyond control of variables: What needs to develop to achieve skilled scientific thinking? *Cognitive Development, 23*(4), 435–451. https://doi.org/10.1016/j.cogdev.2008.09.006

Kuhn, D., & Pearsall, S. (1998). Relations between metastrategic knowledge and strategic performance. *Cognitive Development, 13*(2), 227–247. https://doi.org/10.1016/S0885-2014(98)90040-5

Lawson, A. E. (1978). The acquisition of propostional logic and formal operational schemata during the secondary school years. *Journal of Research in Science Teaching, 15*(6), 465–478. https://doi.org/10.1002/tea.3660150605

Liu, S., & Liu, M. (2020). The impact of learner metacognition and goal orientation on problem-solving in a serious game environment. *Computers in Human Behavior, 102*, 151–165. https://doi.org/10.1016/j.chb.2019.08.021

Lotz, C., Scherer, R., Greiff, S., & Sparfeldt, J. R. (2017). Intelligence in action – Effective strategic behaviors while solving complex problems. *Intelligence, 64*, 98–112. https://doi.org/10.1016/j.intell.2017.08.002

Lotz, C., Sparfeldt, J. R., & Greiff, S. (2016). Complex problem solving in educational contexts–Still something beyond a "good g"? *Intelligence, 59*, 127–138. https://doi.org/10.1016/j.intell.2016.09.001

Luwel, K., Torbeyns, J., & Verschaffel, L. (2003). The relation between metastrategic knowledge, strategy use and task performance: Findings and reflections from a numerosity judgement task. *European Journal of Psychology of Education, 18*(4), 425–447. https://doi.org/10.1007/BF03173245

Mainert, J., Kretzschmar, A., Neubert, J. C., & Greiff, S. (2015). Linking complex problem solving and general mental ability to career advancement: Does a transversal skill reveal incremental predictive validity? *International Journal of Lifelong Education, 34*(4), 393–411. https://doi.org/10.1080/02601370.2015.1060024

Mainert, J., Niepel, C., Murphy, K. R., & Greiff, S. (2019). The incremental contribution of complex problem-solving skills to the prediction of job level, job complexity, and salary. *Journal of Business and Psychology, 34*(6), 825–845. https://doi.org/10.1007/s10869-018-9561-x

Molnár, G., & Csapó, B. (2018). The efficacy and development of students' problem-solving strategies during compulsory schooling: Logfile analyses. *Frontiers in Psychology, 9*, 302. https://doi.org/10.3389/fpsyg.2018.00302

Molnár, G., Greiff, S., & Csapó, B. (2013). Inductive reasoning, domain specific and complex problem solving: Relations and development. *Thinking Skills and Creativity, 9*, 35–45. https://doi.org/10.1016/j.tsc.2013.03.002

Montague, M. (1991). Gifted and learning-disabled gifted students' knowledge and use of mathematical problem-solving strategies. *Journal for the Education of the Gifted, 14*(4), 393–411. https://doi.org/10.1177/016235329101400405

Mund, M., & Nestler, S. (2019). Beyond the cross-lagged panel model: Next-generation statistical tools for analyzing interdependencies across the life course. *Advances in Life Course Research, 41*, 100249. https://doi.org/10.1016/j.alcr.2018.10.002

Mustafić, M., Yu, J., Stadler, M., Vainikainen, M.-P., Bornstein, M. H., Putnick, D. L., & Greiff, S. (2019). Complex problem solving: Profiles and developmental paths revealed via latent transition analysis. *Developmental Psychology, 55*, 2090–2101. https://doi.org/10.1037/dev0000764

Muthén, L. K., & Muthén, B. O. (1998–2021). *Mplus User's Guide. Edition 8.6*. Los Angeles.

Neubert, J. C., Mainert, J., Kretzschmar, A., & Greiff, S. (2015). The assessment of 21st century skills in industrial and organizational psychology: Complex and collaborative problem solving. *Industrial and Organizational Psychology, 8*(2), 238–268. https://doi.org/10.1017/iop.2015.14

Ohtani, K., & Hisasaka, T. (2018). Beyond intelligence: A meta-analytic review of the relationship among metacognition, intelligence, and academic performance. *Metacognition and Learning, 13*(2), 179–212. https://doi.org/10.1007/s11409-018-9183-8

Rudolph, J., Niepel, C., Greiff, S., Goldhammer, F., & Kröner, S. (2017). Metacognitive confidence judgments and their link to complex problem solving. *Intelligence, 63*, 1–8. https://doi.org/10.1016/j.intell.2017.04.005

Schoenfeld, A. H. (2017). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. *Journal of Education, 196*(2), 1–38. https://doi.org/10.1177/002205741619600202

Schoppek, W., & Fischer, A. (2015). CPS-single ability or complex phenomenon? *Frontiers in Psychology, 6*(1669), 1–4. https://doi.org/10.3389/fpsyg.2015.01669

Schoppek, W., & Fischer, A. (2017). Common process demands of two complex dynamic control tasks: Transfer is mediated by comprehensive strategies. *Frontiers in Psychology, 8*, 2145. https://doi.org/10.3389/fpsyg.2017.02145

Schwichow, M., Croker, S., Zimmerman, C., Höffler, T., & Härtig, H. (2016). Teaching the control-of-variables strategy: A meta-analysis. *Developmental Review, 39*, 37–63. https://doi.org/10.1016/j.dr.2015.12.001

Shaharabani, Y. F., & Yarden, A. (2019). Toward narrowing the theory–practice gap: Characterizing evidence from in-service biology teachers' questions asked during an academic course. *International Journal of STEM Education, 6*(1), 1–13. https://doi.org/10.1186/s40594-019-0174-3

Shayer, M. (1976). The pendulum problem. *British Journal of Educational Psychology, 46*(1), 85–87. https://doi.org/10.1111/j.2044-8279.1976.tb02989.x

Sonnleitner, P., Keller, U., Martin, R., & Brunner, M. (2013). Students' complex problem solving abilities: Their structure and relations to reasoning ability and educational success. *Intelligence, 41*(5), 289–305. https://doi.org/10.1016/j.intell.2013.05.002

Stadler, M., Becker, N., Schult, J., Niepel, C., Spinath, F. M., Sparfeldt, J. R., & Greiff, S. (2018). The logic of success: The relation between complex problem-solving skills and university achievement. *Higher Education, 76*(1), 1–15. https://doi.org/10.1007/s10734-017-0189-y

Stadler, M., Fischer, F., & Greiff, S. (2019a). Taking a closer look: An exploratory analysis of successful and unsuccessful strategy use in complex problems. *Frontiers in Psychology, 10*, 777. https://doi.org/10.3389/fpsyg.2019.00777

Stadler, M., Niepel, C., & Greiff, S. (2016). Easily too difficult: Estimating item difficulty in computer simulated microworlds. *Computers in Human Behavior, 65*, 100–106. https://doi.org/10.1016/j.chb.2016.08.025

Stadler, M., Niepel, C., & Greiff, S. (2019b). Differentiating between static and complex problems: A theoretical framework and its empirical validation. *Intelligence, 72*, 1–12. https://doi.org/10.1016/j.intell.2018.11.003

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*(15), 1812–1831. https://doi.org/10.1080/09500693.2018.1504346

Teig, N., Scherer, R., & Kjærnsli, M. (2020). Identifying patterns of students' performance on simulated inquiry tasks using PISA 2015 log-file data. *Journal of Research in Science Teaching, 57*(9), 1400–1429. https://doi.org/10.1002/tea.21657

Vainikainen, M. P. (2014). *Finnish primary school pupils' performance in learning to learn assessments: A longitudinal perspective on educational equity. Report 360*. Unigrafia

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

Veenman, M. V., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction, 14*(1), 89–109. https://doi.org/10.1016/j.learninstruc.2003.10.004

Vestal, S. S., Miller, M., & Browning, L. (2017). Metacognition across the STEM Disciplines. In P. L. Daubenmire (Ed.), *Metacognition in Chemistry Education: Connecting Research and Practice* (pp. 17–30). American Chemical Society. https://doi.org/10.1021/bk-2017-1269.ch002

Wang, M. T., Binning, K. R., Del Toro, J., Qin, X., & Zepeda, C. D. (2021). Skill, thrill, and will: The role of metacognition, interest, and self-control in predicting student engagement in mathematics learning over time. *Child Development*. https://doi.org/10.1111/cdev.13531

Wu, H., & Molnár, G. (2021). Logfile analyses of successful and unsuccessful strategy use in complex problem-solving: a cross-national comparison study. *European Journal of Psychology of Education*, 1-24. https://doi.org/10.1007/s10212-020-00516-y

Wüstenberg, S., Stadler, M., Hautamäki, J., & Greiff, S. (2014). The role of strategy knowledge for the application of strategies in CPS tasks. *Technology, Knowledge and Learning, 19*, 127–146. https://doi.org/10.1007/s10758-014-9222-8

Zhang, X., Hu, B. Y., Ren, L., & Fan, X. (2017). Pathways to reading, mathematics, and science: Examining domain-general correlates in young Chinese children. *Contemporary Educational Psychology, 51*, 366–377. https://doi.org/10.1016/j.cedpsych.2017.09.004

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review, 27*(2), 172–223. https://doi.org/10.1016/j.dr.2006.12.001

Zohar A. (2012). Explicit Teaching of Metastrategic Knowledge: Definitions, Students' Learning, and Teachers' Professional Development. In A. Zohar, & Y. Dori (Eds.), *Metacognition in Science Education. Contemporary Trends and Issues in Science Education* (vol 40). Springer. https://doi.org/10.1007/978-94-007-2132-6_9

Zohar, A., & Barzilai, S. (2013). A review of research on metacognition in science education: Current and future directions. *Studies in Science Education, 49*(2), 121–169. https://doi.org/10.1080/03057267.2013.847261

Zohar, A., & Ben David, A. (2008). Explicit teaching of meta-strategic knowledge in authentic classroom situations. *Metacognition and Learning, 3*(1), 59–82. https://doi.org/10.1007/s11409-007-9019-4

Zumbach, J., Ortler, C., Deibl, I., & Moser, S. (2020). Using prompts to scaffold metacognition in case-based problem solving within the domain of attribution theory. *Journal of Problem-Based Learning, 7*(1), 21–31. https://doi.org/10.24313/jpbl.2020.00206

Authors and Affiliations

Björn Nicolay[1] · Florian Krieger[2] · Matthias Stadler[3] · Mari-Pauliina Vainikainen[4] · Marlit Annalena Lindner[5] · Anne Hansen[6] · Samuel Greiff[1]

Florian Krieger
florian.krieger@tu-dortmund.de

Matthias Stadler
matthias.stadler@psy.lmu.de

Mari-Pauliina Vainikainen
mari-pauliina.vainikainen@tuni.fi

Marlit Annalena Lindner
mlindner@leibniz-ipn.de

Anne Hansen
anne.hansen@uzh.ch

Samuel Greiff
samuel.greiff@uni.lu

1 Faculty of Humanities, Education and Social Sciences, University of Luxembourg, Campus Belval, Maison des Sciences Humaines, 11, Porte des Sciences, L-4366 Esch-sur-Alzette, Luxembourg

2 Faculty of Rehabilitation Sciences, Technical University Dortmund, Emil-Figge-Str. 50, 44227 Dortmund, Germany

3 Faculty of Psychology and Pedagogy, Ludwig-Maximilians-Universität München, Leopoldstraße 13, 80802 Munich, Germany

4 Faculty of Education and Culture, Tampereen yliopisto, Tampere University, Åkerlundinkatu 5, 33100 Tampere, Finland

5 Leibniz-Institut Für Die Pädagogik der Naturwissenschaften Und Mathematik, Olshausenstraße 62, 24118 Kiel, Germany

6 Department of Psychology, University of Zurich, Binzmuehlestrasse 14, Box 1, 8050 Zurich, Switzerland