OULUN YLIOPISTO

Quan Do

# Facets of metacognition and collaborative complex problem-solving performance

Master's thesis

FACULTY OF EDUCATION

Learning, Education and Technology, Master's Programme in Education

15.06.2020

## Abstract:

Metacognition refers to students' ability to reflect upon, understand and control their own learning. Previous accounts of metacognition have distinguished between two major facets: metacognitive knowledge and metacognitive regulation, in which each major facet includes several sub-facets. Although many studies on metacognition facets have examined their relationship with problem-solving performance, few studies have investigated their relationship with non-routine, complex problem-solving performance in collaborative context. In light of this, the current study investigated the impact of different facets of metacognition on perceived and objective complex problem-solving (CPS) task performance in collaborative situation.

Data was collected from 77 students at the University of Oulu, Finland. The Metacognitive Awareness Inventory (MAI) self-report was used to measure subjects' beliefs on the facets of their metacognition before the task. After filling out MAI self-report individually, participants gathered in groups of 3 to carry out the collaborative CPS task. The Tailorshop microworld simulation was employed as the CPS task and used to measure objective group performance. Perceived individual and group performances were measured with self-report. A generalized estimating equation was used to observe the relationships between individuals' awareness of metacognition facets and perceived individual CPS performance. Best Linear Unbiased Predictors (BLUP) function was utilized to yield groups' unbiased MAI scores and unbiased perceived group performance. Pearson correlation coefficient was calculated to observe relationships between group MAI scores and objective group CPS performance, as well as perceived group performance and objective group CPS performance.

In general, the results showed significant correlations between several regulatory facets of metacognition and perceived individual CPS performance as well as objective group CPS performance. Since the majority of the significant correlations were negative, the results reinforced previous findings on students' overconfidence in their skills in relation with their perceived and objective performance as well as contribute to the overall understanding of the impact metacognitive facets have on collaborative CPS performance. Further discussions were addressed in this study. Limitations and future research were also outlined.

**Keywords**: metacognition facets, collaborative problem solving, complex problem solving

## Acknowledgements

I would like to thank supervisor Muhterem Dindar and supervisor Eetu Haataja for guiding me through this thesis work. This is my first time doing a thesis therefore I need a lot of guidance. Thank you for being patient, giving me detailed feedback and encouragement.

Thank you LET (Learning, Education and Technology) programme, faculty of education, university of Oulu for giving me the opportunity to learn, explore learning science and practice scientific thinking, researching and writing skills throughout the 2-year programme.

*Contents*

# 1 Introduction

Every day, students encounter a variety of problems, whether in professional, private, or educational settings. Enhancing students' problem-solving skills is important to ensure their success in life. In education domain, researches have endorsed the importance and relevance of problem solving skill (Greiff, Wüstenberg, et al., 2013b). In one of the most comprehensive international large-scale assessments today, PISA (Programme for International Student Assessment), complex problem solving (CPS) skills were included to be an important measure of competency for dealing with non-routine problems, assessing more than half a million 15-year-old students in over 70 countries (e.g., OECD, 2014).

CPS can be defined as a process of a person acquiring and applying the knowledge regarding the systems that have goal-oriented control and contain many highly dynamic and interrelated elements (Fischer et al., 2017). Studies have showed empirically that people succeeding with problem solving are aware of the gaps between their understanding and the task, and know how to employ problem-solving strategies to close those gaps (Chi et al., 1989). Furthermore, it has been proved that good problem solvers are more active in monitoring their solution progress and reflect more frequently than their counterparts (Chi et al., 1989; Schoenfeld & Herrmann, 1982). This shift from processing the problem itself to monitoring, directing, and evaluating during the problem-solving process in the search for a solution can be described as a shift from cognition to metacognition level (Opwis, 1996).

Besides solving problem independently, students also see themselves in daily situations that require collaboration with others and providing joint efforts and ideas to solve more complex problems. In fact, the concept collaborative problem solving has become prominent recently and was incorporated in PISA assessment in 2015 as one of the most important skills in 21st century (Binkley et al., 2012; Graesser et al., 2017).

Collaborative problem solving can be defined as "the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills and efforts to reach that solution" (OECD, 2017). It mainly involves two facets: problem solving, emphasising on individual cognitive aspects, and collaboration, focusing on the social aspects of teamwork (OECD, 2017; Dillenbourg, 1999; Roschelle & Teasley, 1995). Besides cognitive aspects, participation and interaction are argued to be foundation of group work that leads to successful collaboration (Barron, 2000a, 2003). Studies showed that group having participants monitor their own and their peers' thinking seem to have an advantage over groups who do not (Goos et al., 2002; Hurme et al., 2006). Monitoring function involves assessment of knowledge, strategy, progress toward goal, and result accuracy. This implies that in collaborative problem solving, shared metacognitive events at the social level are more critical than individual metacognitive ones (Malmberg et al., 2017, 2019).

Metacognition can be understood as what we know about our cognitive processes and how we use these processes to solve problem (Kluwe, 1982). While metacognitive knowledge is our understanding and knowledge about our own cognitive activities and processes, metacognitive regulation is what we actually do and engage in to find solutions for a problem. (Brown, 1987; Flavell, 1979; Schraw & Dennison, 1994, Schraw & Moshman, 1995).

It has been showed that good problem solvers are metacognitive problem solvers who direct and regulate their own cognition (Järvelä et al., 2016; Järvenoja et al., 2015; Weinert & Kluwe, 1984). While there are many researches that investigate metacognition and how it relates to problem solving (e.g., Brand et al., 2003; Goos & Galbraith, 1996; Lucangeli et al., 1997; Swanson, 1990), collaborative problem solving (e.g., Goos et al., 2002; Hurme et al., 2006), and task complexity (e.g., Kim et al., 2013), there is a lack of researches examining explicitly the relation between metacognition and collaborative CPS. While the literature supports that metacognitive knowledge and regulation are important for problem solving success, more information is needed to understand how these facets are manifested in a more complex problem-solving context, involves not only independent but collaborative task activities that influence the final performance.

This study aims to investigate the relationship between facets of metacognition and collaborative complex problem-solving performance. 77 participants were involved in this study. Firstly, they filled in the Metacognitive Awareness Inventory (MAI) scale that measured the facets of metacognition (Schraw & Dennison, 1994a) and were randomly grouped into 26 teams to perform the collaborative task. The task is Microworld "*The Tailorshop*" simulation representing a complex problem-solving scenario where participants had to operate a shirt manufacturing company (Danner, Hagemann, Holt, et al., 2011). The correlational analysis between the MAI scale and both individual and collaborative CPS performance was implemented. The study is hoped to broaden understanding regarding the influence of metacognition on Collaborative CPS.

# 2 Theoretical background

This research is centered around the concepts of metacognition and complex problem solving. Collaborative problem solving is presented as a social form of complex problem solving due to a lack of literature on collaborative complex problem solving. Metacognition in complex problem solving are also introduced and discussed.

Collaborative problem solving involves group members combining their resources, skills, and efforts to solve problems and reach desired goals (OECD, 2013). The two main components of collaborative problem solving have been defined as cognitive and social (Hesse, Care, Buder, Sassenberg, & Griffin, 2015). Research has shown that it is challenging to collaborate successfully and requires the effective coordination of individual and group processes in both the cognitive and social dimensions (Barron, 2003; Hadwin et al., 2017). In this sense, collaborative problem solving involves high levels of metacognition to control those processes (Dierdorff & Ellington, 2012; Flavell, 1979).

## 2.1 Metacognition

In this chapter, the researcher addresses the issue of metacognition first by reviewing its definition, historical development, findings and implications. I then provide definition and explanation for each facet of metacognition together with its sub-components. In particular, calibration of performance in relation to metacognitive monitoring is also explained. Throughout, the primary emphasis is around the concept of metacognition and its components as variables being investigated in this thesis.

### 2.1.1 Metacognition definition, historical overview, findings and implications

The term metacognition was originally coined by John Flavell in the late 1970s to mean “cognition about cognitive phenomena”, or “thinking about thinking” (Flavell, 1979). In education, it often refers to students’ ability to reflect upon, understand, and control their own learning (Brown, 1987; Schraw & Dennison, 1994).

Despite different operationalizations from different fields, most researchers agree that metacognition has mainly two functions: monitoring and controlling of cognition (Dunlosky & Metcalfe, 2008). In a general framework describing the relationship between metacognition and cognition, Nelson (1984) refers cognition to as the object level, in which the cognitive processes such as learning, problem solving take place. Metacognition is referred to as the meta-level representing our understanding of the ongoing cognitive processes that we are engaged while performing the task. In terms of this framework, at the meta-level, metacognition plays a role of monitoring the current state of the object-level, meaning monitoring the cognitive activities and processes taking place. After receiving information from the object-level, metacognition can control those cognitive processes (e.g., try different strategies) to achieve the learning goal or increase the task performance (Dunlosky & Metcalfe, 2008). Those monitor and control function are closely related because whether or not we engage in metacognitive control, will depend on metacognitive monitoring's outcome. Similarly, metacognitive monitoring would only be useful if it could trigger metacognitive control when needed (Nelson, 1984) .

In psychology, metacognition has been studied across disciplines from cognitive psychology, personality psychology, social psychology, to more applied settings such as clinical psychology and educational psychology (Elisabeth Norman et al., 2019). It has become even more prominent since 1979, when Flavell proposed a framework to distinguish different facets of the phenomena (Flavell, 1979). Contributions of researches on metacognition in educational psychology have mainly centered around two major themes: (a) Possibility and methodology of teaching metacognitive skills, and (b) Influence of metacognition on learning (Elisabeth Norman et al., 2019). Regarding the former theme, studies have showed that efforts on teaching metacognitive skills in schools have made a positive impact on students learning. This was convinced evidently through meta-analysis of Donker et al., (2014) and Dignath et al. (2008). Similarly, there has been a large volume of research examining the latter theme, the link between metacognition and learning, through a measurement of academic achievement. However, in a meta-analysis of (Dent & Koenka, 2016) related to the theme, the association seems to be relatively weak (mean, weighted correlation of r=.20). In which, the association was stronger when measured using online self-report comparing to the offline one. Thus, there is a need of conducting more researches to examine the influence of metacognition on learning.

### 2.1.2 Facets of metacognition

Researchers have tried to distinguish different sub-components of metacognition. Previous accounts of metacognition have distinguished between two major components: knowledge about cognition and regulation of cognition (Brown, 1987; Flavell, 1979; Schraw & Dennison, 1994a). In this section, I elaborate on the distinction between metacognitive knowledge and metacognitive regulation and describe subcomponents involved in each.

*Knowledge about Cognition (also referred as Metacognitive Knowledge)*

According to Schraw & Moshman (1995), "*knowledge of cognition refers to what individuals know about their own cognition or about cognition in general. It usually includes three different kinds of metacognitive awareness: declarative, procedural, and conditional knowledge* (Brown, 1987; Jacobs & Paris, 1987)*"*. The distinction between these three kinds of knowledge about cognition is describe as follows:

*Declarative knowledge*. Declarative knowledge includes what students know about themselves and different aspects that influence their learning performance (Schraw & Moshman, 1995). For instance, good learners have been showed to have more knowledge about their own memory and its usability comparing to poor learners (Garner, 1987; Schneider & Pressley, 1989). Similarly, according to Jacobs and Paris (1987), student's prior knowledge and familiarity with a reading topic play an influential role in their reading speed and comprehension. Comparing with children, researches on metamemory have showed that adults appear to possess more knowledge about their cognitive processes associated with memory (Baker, 1989). In recent years, declarative knowledge has been expanded to include individuals' knowledge and understanding of their affective states, including self-efficacy and motivation, and how these aspects affect task performance (Pressley & Gaskins, 2006).

*Procedural Knowledge*. Procedural knowledge refers to knowledge about the execution of thinking processes. Students with high procedural knowledge could employ different strategies and skills (e.g., skimming, underlining, summarising while reading) to complete a task (Schraw & Moshman, 1995). The higher degree of students' procedural knowledge, the more automatically, effectively and qualitatively they can use and sequence different strategies to solve problem (Glaser et al., 1988; Stanovich, 1990). Studies have showed that students who received procedural training perform their on-line problem-solving better (King, 1991). The acquisition of procedural knowledge has been proved to influence tasks' concepts acquisition as a result increase task understanding and performance (Engelbrecht et al., 2005).

*Conditional Knowledge*. Conditional knowledge refers to our knowledge of the conditions that influence learning. In particular, it refers to knowing why various strategies are effective and when they are applicable and appropriate (Garner, 1990; Lorch et al., 1993). For example, college students have been showed to employ very different strategies from each other to process a reading task; in which they selected the strategies that are most appropriate for their own situation (Lorch et al., 1993). In a review of (Reynolds, 1992), older children and adults appear better at selectively allocate their attention based on the condition of the task at hand comparing to younger learners. Furthermore, (Justice & Weaver-McDougall, 1989) have found in adults that their knowledge of various strategies use (conditional knowledge) has a positive correlation with the act of using the strategies themselves (regulation of cognition). Comparing to individuals who have lower conditional knowledge, individuals possessing high degree of conditional knowledge can better assess and estimate the demands of their specific learning situation, select most appropriate strategies and procedure for that situation. (Schraw, 2009).

In sum, researches have claimed that good learners appear to have knowledge of cognition including declarative, procedural and conditional knowledge about cognition and the possession of this knowledge helps them perform better. In comparison, older learners have the tendency to possess more knowledge about their cognition. Interestingly, although possessing metacognitive knowledge or knowledge of cognition, learners appear to find it very difficult to explicitly describe it (Bereiter & Scardamalia, 1993).

*Regulation of Cognition (also referred as Metacognitive Regulation)*

Regulation of cognition, or metacognitive regulation, refers to the activities that we do and engage in to control our thinking and learning. It includes five sub-components that facilitate this control aspects: planning, information management strategies, comprehension monitoring, debugging strategies, and evaluation (Artz & Armour-Thomas, 1992; Baker, 1989; Schraw & Moshman, 1995)

*Planning*. Planning involves the activities of setting goals, planning and allocating resources in prior to learning. For examples, before a writing task, we set a goal of how many pages we want to write, then allocate time and our attention exclusively before we start (Miller, 1985). Studies about this metacognitive regulatory subcomponent have found that more experienced learners are more effective in planning regardless of the learning content. Comparing to less experienced learners, more experienced ones have more knowledge about their cognition and know how to utilise them to regulate their learning in prior to the task performing phase (Baker, 1989; Bereiter & Scardamalia, 1993; Garner & Alexander, 1989). Moreover, planning ability has been showed to exist in young learners and develop significantly in teenage years (Bereiter & Scardamalia, 1987). Comparing to metacognitive monitoring and evaluation, planning tended to be applied more by individuals and mature earlier (Veenman et al., 2006). According to (Schraw et al., 2006), planning in prior to task activity may improve outcome regardless of context and task content.

*Information management strategies.* Information management strategies refer to strategy sequences used to process information more efficiently. Some examples can be organizing, summarising information, elaborating a concept for further understanding, or focusing selectively on important information. Studies of (Smith, 1982, 1985) and (Bereiter & Bird, 1985) on graduate students and professionals, who were considered "expert" readers reported that good readers employed a variety of information management strategies such as problem formulation, rephrasing the materials into simpler terms, focusing on elaborating important information. In comparison between good and poor readers, studies suggested that good readers reported using more strategies when they encountered a difficulty in comprehension (Kaufman et al., 1985) and more especially, they were more persistent in their efforts to understand and finish the reading task comparing to poor readers (Kletzien, 1988; Spring, 1985).

*Comprehension Monitoring*. Comprehension Monitoring refers to our assessment of our learning, thinking processes or strategy use (Cera et al., 2013; Schraw & Moshman, 1995). For example, during a learning task, we employ comprehension monitoring by doing regular self check-in and testing on our on-going performance. Most of the studies in comprehension monitoring have reported that subjects regardless of ages, are quite poor in monitoring their comprehension and even overconfident in their performance (Glenberg & Epstein, 1987; Pressley & Ghatala, 1990). However, studies also indicate that monitoring ability can be trained. In a study of (Delclos & Harrington, 1991), young students engaging in a computer problem-solving task, who were trained how to monitor their performance appeared to solve more difficult problems and perform faster comparing to non-trained ones. Monitoring is important because it can lead to improved understanding of content and problem-solving ability (Serra & Metcalfe, 2009).

*Debugging strategies.* Debugging strategies refer to a variety of "fixup" strategies to correct comprehension and performance errors. Some examples can be when facing confusion, we stop and reread previous text or consult outside resources such as books and other people. Studies showed that students spontaneously use rereading strategy upon encountering a sentence that conflicts with previous information. And by using rereading strategy, students reported improved test performance (Baker & Anderson, 1982). Additionally, students who were induced to consult outside resource such as looking up information that they didn't understand subsequently recalled more information than those without the option (Blohm, 1987).

*Evaluation*. Evaluation refers to the appraisal of the learning outcomes and the effectiveness of our regulatory processes after the learning task. For example, after a learning episode, we re-evaluate our learning outcomes and whether they have met our learning goals. Studies have showed that learners who possess knowledge about their cognition and planning skills tend to have good evaluation of their learning (Baker, 1989). In their research, (Bereiter & Scardamalia, 1987) showed that good writers had the tendency to adopt more objective goal and evaluation of their work and try to revise their work to the readers' point of view while poor writers had more difficulties in adopting and correcting them. Evaluation has been found to be interconnected with planning and conditional knowledge in which when subjects evaluate their performance, they might consider planning differently in relation with the strategies use and conditional factors affected their performance (Tanner, 2012).

Overall, researchers agree that possessing metacognitive regulation improves learner's performance. It can help them utilise cognitive resources such as attention more effectively, employ various and better use of strategies to process information, fix errors, and maintain a greater awareness of their comprehension and performance. Researches also indicate that the two sub-components, knowledge of cognition and regulation of cognition are not independent but inter-related with each other (Brown, 1987; Flavell, 1979; Schraw & Dennison, 1994a).

*Metacognition and Calibration of Performance*

Besides helping manage attentional resource or process information more effectively, metacognition also plays an important role in helping problem solver monitor their performance more accurately (Schraw et al., 2000). Monitoring falls under the regulation facet of metacognition. It refers to subjects' awareness of comprehension and task performance while doing a task (Baker & Brown, 1984; Flavell, 1979; Nietfeld et al., 2005; Schraw, 1995). Monitoring is important because it helps generate feedback for a person to control their task and use regulatory strategies to solve problem and improve their performance (Ellis & Zimmerman, 2001). Confidence judgement is one of the methods to operationalize monitoring accuracy (Schraw et al., 2000). It can then be analysed to arrive at either relative or absolute measures of monitoring accuracy. An arrival at a measure of absolute accuracy is termed calibration (Nelson, 1984; Schraw, 1995).

According to Nietfeld et al. (2005), in general, calibration is the process of matching perception of performance with actual level of performance. There are typically two types of calibration: calibration of comprehension and calibration of performance (Glenberg & Epstein, 1987). Calibration of comprehension refers to students' confidence judgement in their ability to acquire knowledge or perform a forthcoming task. Calibration of performance refers to students' confidence judgement in their output or produced answer after completing the task (García et al., 2016; Nietfeld et al., 2005). Many studies on calibration have showed that calibration of performance tends to be more accurate than calibration of comprehension (Glenberg & Epstein, 1985, 1987; Maki et al., 1990). One of the possible reasons is due to the additional feedback students gain while performing the task (Ackerman & Wolman, 2007).

Accurate calibration, in relation with metacognitive monitoring accuracy, plays a crucial role in alerting students about their ongoing cognitive processes and using regulatory strategies to efficiently control their problem-solving performance (Ellis & Zimmerman, 2001; Everson & Tobias, 2001; Lin et al., 2002). The more students' judgement of confidence matches their actual performance, the better calibrated the students are (Hacker et al., 2008). In this sense, students' ability to accurately gauge their progress and performance plays an important role in their following effort and strategy use in problem-solving situations (Alexander, 2013). However, accurate calibration is not common for both students and adults. Low-performing students prevalently show overconfidence when performing a task (Bouffard et al., 2011; Dinsmore & Parkinson, 2013; Hadwin & Webster, 2013). Furthermore, many studies have showed that calibration judgements have the tendency to become stable and resistant to improvement over time (Bouffard et al., 2011; Hacker et al., 2008).

There are different methods to calculate calibration. Among them are dichotomous ratings and categorical ratings (Parkinson et al., 2010). In dichotomous ratings, students complete multiple-choice of Confident or Not Confident about each answered item of the task (Schraw et al., 2012). In categorical ratings, such as Likert-type scales, students measure their confidence by rating a scale ranging from not confident to very confident (Hattie, 2013). Judgement accuracy then expressed by different measures such as Pearson correlations or Gamma coefficient (Dinsmore & Parkinson, 2013; Schraw, 2009). In the present study, a categorial measure of calibration was used.

In sum, this section elaborates and explains the concepts of metacognition, its two facets metacognitive knowledge and metacognitive regulation as well as eight sub-components under the two facets. Additionally, calibration of performance was also addressed as a part of metacognitive regulatory monitoring dimension. These concepts related to metacognition have been showed to appear in good learners and help learners to improve their performance. However, the studies examining these concepts although are foundational, they didn't address directly the relationship with collaborative CPS which proves to be important and relevant in the modern age. Therefore, there is a need to provide more present studies investigating metacognition and how it influences complex problem-solving performance in group contexts.

## 2.2 Complex Problem Solving (CPS)

According to Fischer et al. (2017), CPS can be formally defined as a process of a person acquiring and applying the knowledge regarding the systems that have goal-oriented control and contain many highly interrelated elements. Dörner & Funke (2017) later redefined CPS to make it less formal and close to the real world. The new definition emphasised that "*Complex problem solving is a collection of self-regulated psychological processes and activities necessary in dynamic environments to achieve ill-defined goals that cannot be reached by routine actions. Creative combinations of knowledge and a broad set of strategies are needed. Solutions are often more bricolage than perfect or optimal. The problem-solving process combines cognitive, emotional, and motivational aspects, particularly in high-stakes situations. Complex problems usually involve knowledge-rich requirements and collaboration among different persons.*"

The concept was first introduced in Germany by Dörner and colleagues in the 1970s (Dörner, 1975; Dörner et al., 1975). The need for researches on CPS has been rising ever since. Some examples of recent complex problems are about global climate politics (e.g. Amelung & Funke, 2013), mismanaged corporations (e.g. Funke, 2003), or nuclear disasters (e.g. Dörner, 1997). Today, CPS skills proved to be an important measure of competency in solving non-routine problems in large-scale assessments such as PISA ("Programme for International Student Assessment") (OECD, 2014) and is recognized as one of the most important competencies required in the future (World Economic Forum, 2015).

According to Funke (2010), a complex problem has following characteristics: (a) situation complexity, involving a large number of variables; (b) mutual dependencies and connectivity between those variables; (c) dynamic development of the situation over time; (d) opaqueness of the variables and their current values; and (e) representing goal conflicts on different levels of analysis.

### **2.2.1** Assessment of CPS

The CPS competency is presented in the form of operative intelligence, a concept proposed by Dörner (1986). Operative intelligence means a person's ability to coordinate his or her cognitive operations in the pursuit of a well-informed and sustainable goal. It involves knowledge of declaration, procedure, and contexts in which the operations play out (Fischer, Greiff, Funke, 2017). Because the nature of CPS involves many processes going on at the same time, and no such interaction is possible within paper-pencil tests (Matthias Stadler et al., 2015). As a result, researchers have employed computer-based CPS simulations to assess operative intelligence. These computer-based CPS simulations can be derived from the states or processes of the system's variables, inputs of problem solvers, or knowledge tests conducted during or after the process (Dörner, 1986a; Greiff, 2012; Strohschneider & güß, 1999).

There is a great variety of assessment approaches used to measure CPS. The first computer-based CPS tasks were designed to resemble the real world, with the aim of providing ecological valid measure of CPS. Some examples can be leading a business by manipulating 24 variables for 12 simulated months using microworld Tailorshop simulation (Danner et al., 2011); or governing a city through microworld Lohausen simulation which involves 1000 interrelated variables (Dörner, Kreuzig, Reither, & Stäudel, 1983). Although these classical CPS measures have high ecological validity, there were arguments that they had shortcomings related to psychometric properties (Greiff, Stadler, Sonnleitner, Wolff, & Martin, 2015).

Joachim Funke (2001) introduced two formal frameworks to describe fundamental structures of the CPS tasks, called Linear Structural Equation systems (LSE) and Finite State Automata (FSA). Many CPS researches have adopted LSE for the development of single complex systems (Kröner, 2001; Wagener, 2001) and later advanced to multiple complex systems (Greiff, Fischer, et al., 2015). Since employing multiple small microworlds as measures of CPS, with the focus on psychometric quality, multiple complex systems show significantly higher reliability of measurement comparing to classical measures of CPS which (Funke, 2001) use one single, large microworld. Assessment tools utilising multiple complex systems approach can be named such as MicroDYN (Greiff et al., 2012), Genetics Lab (Sonnleitner et al., 2012), and MicroFIN (Neubert et al., 2014).

While there are different assessment approaches of measuring CPS with their own pros and cons in reliability and validity, Joachim Funke et al. (2018a) suggested that new methods of CPS assessment should meet the construct validity of the classical microworlds but have the psychometric properties of the multiple complex systems. In the context of this research, classical microworld Tailorshop is employed to investigate individuals complex problem-solving performance.

### **2.2.2** Researches on CPS

Researches have showed a wide range of aspects that influence CPS performance. From task properties such as time (e.g., Greiff et al., 2016), to participants behaviours such as planning behaviour (Eichmann et al., 2019), participants' GPA (e.g, Greiff, Wüstenberg, et al., 2013), and their motivation (e.g., Vollmeyer & Rheinberg, 1999, 2000). Besides these aspects, majority of the CPS researches is on cognitive aspects and processes (Güss et al., 2017) in which metacognitive aspects and processes have also been investigated by a few empirical studies (e.g., Rudolph et al., 2017; Toy, 2007; Shin et al., 2003).

Regarding cognitive aspects, a study of Wüstenberg et al. (2012), employed Advanced Progressive Matrices measurement and the MicroDYN CPS task, had found that reasoning was a main predictor for both facets of CPS, namely knowledge acquisition ($\beta = 0.63$, $p < 0.001$, $R^2 = 0.39$), and knowledge application ($\beta = 0.56$, $p < 0.001$, $R^2 = 0.31$). Additionally, Greiff, Fischer, et al. (2013) shared the similar results and even found that CPS correlated with not only reasoning but other cognitive abilities which supported previous studies (Danner, Hagemann, Holt, et al., 2011; Greiff, Wüstenberg, et al., 2013a).

Besides cognitive abilities, researchers also investigated the importance of CPS knowledge in CPS performance. Knowledge demanded in a CPS task can include structural knowledge, usually referred to CPS knowledge acquisition facet, and strategic knowledge, referred to CPS knowledge application (J. Funke, 1991, 2012; Fischer et al., 2012; J. Funke, 1991). For instance, Funke (1988) utilised ECOSYSTEM CPS task to investigate knowledge acquisition, knowledge application. The findings indicated that the development and application of structural knowledge of and about the system is a must to describe the relationships and interconnectedness of the variables in the system.

Besides the influence of cognitive processes, success of CPS is believed to also require regulation of cognition (Dörner & Funke, 2017a; Joachim Funke, 2010). With the aim to better understand the nature of CPS, Rudolph and colleagues (2017) investigated the link between confidence judgments, which are main presentation of metacognitive self-monitoring, and CPS confidence, with control of reasoning. They used judgements of confidence after each phase of MicroDYN task to assess monitoring of knowledge acquisition and knowledge application. The results suggested that confidence in CPS task explained 55% of the variance for knowledge acquisition and 68% for knowledge application ($p<.001$). These findings indicate that confidence judgments as indicators of metacognitive monitoring in CPS are substantially linked to successful CPS.

Another empirical study of Shin et al. (2003) showed a similar relationship. In their study, problem-solving skills were compared in solving well-structured problems and ill-structured problems. Instruments were developed and administered to assess cognitive and affective predictors of problem-solving performance. The results showed that domain knowledge and justification skills significantly predicted well-structured problem-solving performance, whereas domain knowledge, justification skills, science attitudes, and regulation of cognition were significant predictors of ill-structured problem-solving performance. Specifically, metacognitive regulation was found to be positive correlated with complex problem-solving performance.

However, another study of Toy (2007) showed a different relationship. The study investigated the relationship between metacognitive strategies used by university students and task performance. The task was defined as complex, ill-structured problem-solving task in web-based environment. The results showed that the metacognitive strategies students used, measured by MAI, were significant but negative predictors of problem-solving performance when learner characteristics had been controlled. Specifically, regulation of cognition components were negatively correlated with students' scores on complex, ill-structured problem solving task performances. While metacognitive awareness is hypothesized to enhance performance in solving problems that are complex and ill-structured (Flavell, 1979; Jacobs & Paris, 1987), these studies have presented contradictory results on the role of metacognition on complex problem-solving performance.

In conclusion, in this section, firstly we have discussed definitions of CPS, with a stress on knowledge acquisition and knowledge application processes in a highly dependent and interrelated system of variables. Secondly, we have analysed different approaches to assess CPS, namely classical microworlds and multiple complex systems. Advantages and disadvantages of both approaches have been discussed which justifying our selection of Tailorshop as a CPS assessment in this research. Thirdly, we investigated different studies on CPS. CPS task is influenced by a wide range of aspects, in which cognitive aspects appear as a main topic. Especially, metacognition, a key concept in this research, was also referred and showed to be correlated with CPS performance.

## 2.3 Collaborative Problem Solving

Collaborative problem solving can be defined as "*the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills and efforts to reach that solution*" (OECD, 2017). Collaborative problem solving has become a prominent concept recently. It was incorporated in PISA assessment in 2015 and considered as one of the most important skills in 21st century (Binkley et al., 2012; Graesser et al., 2017). Although collaborative problem solving is complex, it mainly involves two facets: problem solving and collaboration (OECD, 2017). If problem solving emphasises on cognitive aspects of individuals, collaborative facet focuses on the social aspects of teamwork, including coordinating the tasks and maintaining joint problem space (Dillenbourg, 1999; Roschelle & Teasley, 1995). In education, collaboration is often studied in the form of collaborative learning which can be defined as a learning process where individuals actively participate and strive to achieve a common goal (Roschelle & Teasley, 1995). Collaborative learning is grounded in socio-constructivist approach, in which knowledge is shaped and emergent in the social context. Learning mechanisms are triggered when group members continuously and effortfully negotiate, share and discuss meanings (O'Donnell & Hmelo-Silver, 2013).

Participation and interaction in are argued to be foundation of group work that leads to successful collaboration (Barron, 2000a, 2003). A quality group interaction often involves team cognition, socio-emotional factors, reciprocity, and team regulation (Hung, 2013; Isohätälä et al., 2018; Volet, Vauras, Salo and Khosa, 2017; Järvelä, Järvenoja, Malmberg, & Hadwin, 2013).

Firstly, team cognition originates from the interdependence of team process and behaviours in which there is a blurred boundary between individual cognition (Cooke, Salas, Kiekel, & Bell, 2004). Cognition of the team is presented as more than the aggregate or collection of individual cognitions (Hung, 2013; Stahl, Koschmann, & Suthers, 2006) but also influenced by the social context of team members working together as a whole (Hung, 2013). According to Van de Ven et al. (1976) and Wageman (1995), teams that deal with complex problems work within intensive interdependence, which requires greater coordination patterns compared to lower levels of interdependence. Team members must simultaneously diagnose, problem-solve, and coordinate as a team to accomplish a task. Therefore, complex problem-solving situations in teams not only require team members to process information individually (e.g., encoding, storing, retrieving, representing, anticipating), but also demand interdependent actions of all team members for effectively using all resources to reach high team performance (Isohätälä et al., 2017; Järvelä et al., 2013; Järvelä, Kirschner, et al., 2016; Malmberg et al., 2015, 2017).

Secondly, productive group interactions also deal with socio-emotional aspects which often related to group communication, in particular is group argumentation (Isohätälä et al., 2018; Linnenbrink-Garcia et al., 2011). For instance, in a research conducted by Isohätälä et al. (2018) to investigate the role of argumentation in collaborative tasks, their results showed that apart from high-levels of critical cognitive processing, argumentation may also raise heated socio-emotional situations. Similarly, the study conducted by Linnenbrink-Garcia and colleagues (2011) found that negative socio-emotional interactions hinder collaboration while groups with positive interactions were more sustainable.

Thirdly, studies also stress on teams' regulation as an essential key for successful collaborations (Järvelä et al., 2008; Järvelä, Järvenoja, Malmberg, & Hadwin, 2013). According to Järvelä and colleagues (2013), shared regulation refers to processes by which group members actively and collaboratively regulate their cognitions, emotions and motivations, resulting in socially and individually self-regulated members. The regulation can be between peers (co-regulation), and learning collectively in groups (socially shared regulation) that contributes to reciprocal collaborative interaction (Järvelä et al. 2016) and progress in their collaborative learning (Malmberg et al. 2017). Team regulation is emphasised to be situated, depending on the context and the environment such as tasks, methods of intervention, and participants' traits. In the study conducted by Järvelä and colleagues (2008) it was found that face-to-face collaborative tasks triggered emotion and motivation regulation processes at the social level as a result of the group's shared regulation.

## 2.4 Metacognition and CPS

Due to a lack of researches examining explicitly the relation between metacognition and CPS, this section reviews briefly researches that investigate metacognition and how it relates to problem solving (e.g., Aljaberi & Gheith, 2015; Brand et al., 2003), collaborative problem solving (e.g., de Bruin & van Gog, 2012; Goos et al., 2002; Hurme et al., 2006; Salonen et al., 2005), and task complexity (e.g., Kim et al., 2013).

Regarding problem-solving performance, researchers have consistently found that it is influenced by independent metacognitive factors that overlap and interact in a variety of ways. Metacognitive knowledge components, including declarative knowledge, procedural knowledge, and conditional knowledge were found to be most influential (Carlson & Bloom, 2005). Researchers also points out that the processes of control, including metacognitive monitoring and planning are one of the most critical metacognitive dimensions which largely influence the process of decision making in problem solving (Wilson & Clarke, 2004).

In the research of Brand and colleagues (2003), they investigated the influence of metacognitive thinking and knowledge acquisition in improving problem-solving performance. The problems to be solved were several Tower of Hanoi. The participants were formed in pairs (dyads) and stimulated to metacognitive thinking while solving the tasks. The result showed that metacognitive stimulation enhanced performance in all cases. Those participants who had been stimulated to metacognitive thinking, whether individually or in pairs, performed better. Additionally, the dyads showed better problem-solving performance than the individuals.

A study of Aljaberi & Gheith (2015) investigated the relationship between the university students' metacognition and their mathematical and scientific problem-solving performance. The results showed that on the overall scale, metacognition showed no significant correlation with both mathematical and scientific problem-solving performance. However, on mathematical problem-solving alone, it was found to be significantly correlated with metacognitive factors, namely Procedural Knowledge, Evaluation, Debugging Strategies, and Information Management. Additionally, metacognitive regulatory Debugging Strategies component was found to be significantly correlated with both mathematical and scientific problem-solving performance.

In collaborative problem-solving, Goos et al. (2002) and Hurme et al. (2006) also investigated the influence of metacognition. They found that groups having participants monitor their own and their peers' thinking seem to have an advantage over groups who do not. For instance, (Goos et al., 2002) utilised metacognitive monitoring and regulation functions to analyse verbal protocols in collaborative problem-solving setting. Monitoring function involves assessment of knowledge, strategy, progress toward goal, and result accuracy. Regulatory function involves identifying new information, changing strategy, correcting error and so on. As a result, they have found the differences in terms of metacognitive transactive discussions. In poor problem-solving, there was lack of critical engagement in monitoring each other's thinking. This implies that in collaborative problem solving, shared metacognitive events at the social level are more critical than individual metacognitive ones. Similar to this finding, de Bruin & van Gog (2012) and Ucan & Webb (2015) also suggested that monitoring increases progress in collaborative problem solving by enabling group members' reflection on their metacognitive thoughts and feelings about the group's joint progress. Regarding interaction aspect of collaborative problem solving, Salonen et al. (2005) investigated the social interaction in relation with metacognition in a situation involving interaction of students and peers. The study suggested that metacognition plays an important role in enhancing participants' awareness in their own and their peers' cognition and therefore influence metacommunication control processes in collaborative learning.

In a similar vein, to investigate the impact of collaboration in relation with metacognitive monitoring and performance, a similar study of Pao (2014) was also conducted. Participants were randomly assigned to either individual conditions, working alone, or collaborative conditions, discussed the questions with a partner, to answer practice test questions. The results showed that subjects in both conditions were overconfident in their metacognitive monitoring skills. However, peer collaboration was surprisingly found to lead to greater overconfidence comparing to individual condition group. In other words, peer collaboration did not appear to improve metacognitive monitoring but instead became counterproductive in inflating participants' confidence.

Besides problem solving at individual and social levels, several research studies have identified task difficulty in terms of task complexity as an important factor in the elicitation of metacognition (Efklides, 2006; Helms-Lorenz & Jacobse, 2008; Iiskala et al., 2004; Prins et al., 2006; Vauras et al., 2003). Their findings suggest that metacognition has tendency to manifest itself more frequently in complex versus simple tasks. Problems requiring students to deal not only with vague definitions but also individual and social negotiations to construct meaning, share different perspectives, promote multiple solutions to meet commonly agreed team goals, turns out helping elicit metacognitive processes. Thus, a more complex problem seems to evoke more complex metacognitive functions (Kim et al., 2013).

Overall, there is a great confirmation by researchers that failing and succeeding in problem-solving depend on the variety of metacognitive components, from knowledge of cognition to regulatory aspects. Although many studies have investigated and compared characteristics of poor and good problem solvers, in individuals and in groups, the complexity aspect of the problem-solving process still does not appear to be understood. While the literature supports that metacognitive knowledge and regulation are important for problem solving success, more information is needed to understand how these facets are manifested in a more complex problem-solving context, involves not only independent but collaborative task activities that influence the final performance.

# 3 Aims and Research questions

The main aim of this study is to investigate the relationship between facets of metacognition and collaborative complex problem-solving task performance. The research questions (RQ) for this study are:

***RQ1***: *Is there any relationship between facets of metacognition and perceived individual performance in collaborative complex problem solving?*
***RQ2***: *Is there any relationship between facets of metacognition and objective group performance in collaborative complex problem solving?*
***RQ3***: *How accurate is student's calibration of performance in collaborative complex problem solving? The relationship between perceived group performance and objective group performance*

# 4 Methods

## 4.1 Sample

The participants in this study were 77 students at university of Oulu, Finland. The mean age of the participants was 27.8 (SD = 5.5). 53% of the participants were females (f = 41) and 47% were males (f = 33). The majority of the participants were master's graduates (f = 55), followed by PhD students (f = 15) and undergraduates (f = 7). Students' field of studies are as follows: education (36.5%), business, economics or management (33.8%), environmental engineering (5.4%), wireless communication engineering (4.1%), English philology (2.7%), computer science (2.4%), biomedical engineering (1.4%), information processing (1.4%), pharmacology (1.4%), mining (1.4%), engineering (1.4%), architecture (1.4%), ecology (1.4%), product development (1.4%), industrial engineering (1.4%) and not specified (2.4%).

## 4.2 Research design

The researcher employed quantitative methodology for analyzing data in group settings. The design was experimental with no groups control and had the same treatment for all the groups and individuals (Cress & Hesse, 2013). Quantitative research explains phenomena by collecting numerical data that are analysed using mathematically based methods (in particular statistics) (Creswell, 1994). In order to explain the phenomena, the researcher asks the specific questions (e.g., Is there a relationship...?) and collects numerical data from participants to answer the question. The research is correlational, which attempts to determine ″Is there...″ and ″how much ″ a relationship exists between two or variables. The relationship is expressed by correlation coefficient, measured between 00 and 1.00. The numbers are hoped to yield unbiased results that can be generalized to a larger population (Williams, 2007).

## 4.3 The task

The *Tailorshop* is a computer-based dynamic decision-making task (Danner, Hagemann, Holt, et al., 2011). The scenario simulates a small business that produces and sells shirts. The participants are asked to lead the business for 12 simulated months with the aim of maximizing the company value. They do that by manipulating different variables such as increasing or decreasing number of workers, advertising expenses or number of machines (see Figure 1). The Tailorshop consists of 24 variables of which 21 variables are visible to the participants. There are 12 variables which can be manipulated directly (e.g., salary, repair & service costs), whereas the other 12 variables can only be indirectly operated (e.g., demand, worker satisfaction %). The state of a variable is influenced by the state of its own or other variables in every month.

The task is structured in two main phases: exploration and performance. During the exploration phase, the participants can test running the company for 6 simulated months to get familiar with the scenario and the task mechanisms. In every month, participants can interact with the variables by modifying its value, learn more about the variable description by clicking 'i' button on each variable representation for more information. Once they finish a current month and click next to transition to the next month, participants will see the impact of their modification on the other variables so that they can interpret their performance and make change accordingly. Once finishing the exploration phase, participants proceed with performance phase which consists of 12 simulated months. In the performance phase, the simulation starts from the beginning. Therefore, the inputs made in the simulation during exploration has no effect on performance phase.

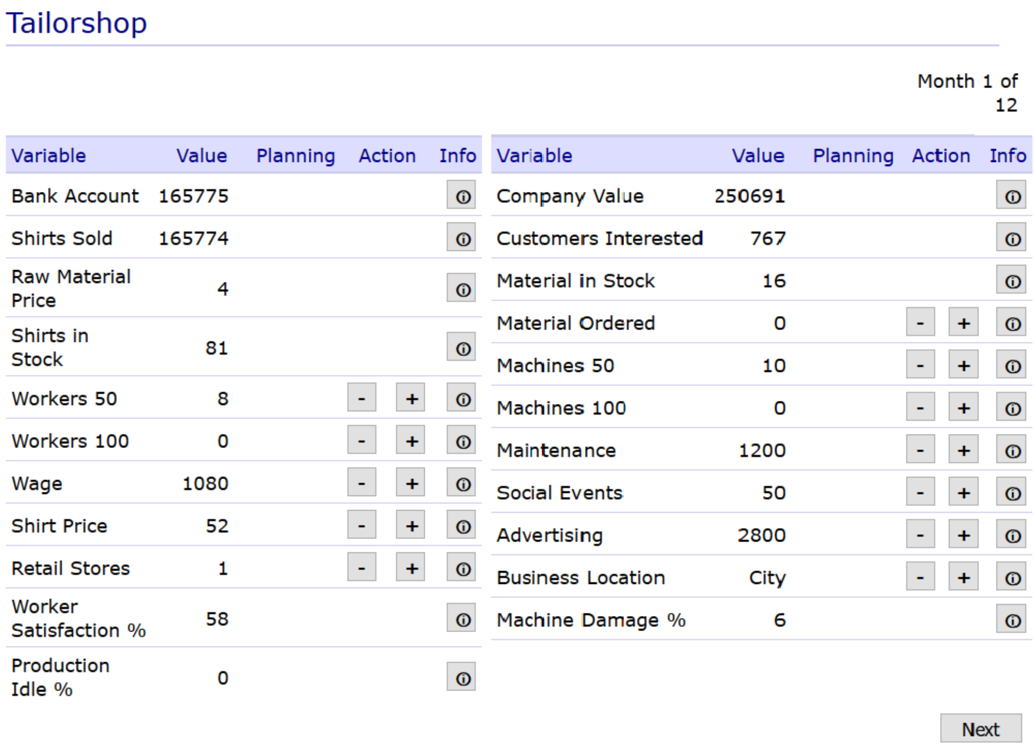


***Figure 1. Screenshot of the graphical user interface of the Tailorshop***

## 4.4 Procedure

Call for research volunteer was carried out at the opening of the academic year at University of Oulu. Participants were reached out through their registration via both online (media announcement, billboard) and offline announcements (distributing handouts at the new academic year opening) at the university campus. It was voluntary to participate in the study. Registration was done through a web link and after that an appointment was made according to participants' availability. Research meetings were conducted at LeaForum (Learning and Interaction Observation Forum) at University of Oulu. The room was well-facilitated for the current study with video tracking, biosensor devices and could host 3 groups at a time.

At the research meeting, participants were asked to sign the consent forms and then proceeded to complete the questionnaires consisting of demographics survey and Metacognitive Awareness Inventory (MAI) scale. They were then invited to join their group of three which was randomly assigned and also based on their availability. Each group was presented with a computer having *Tailorshop* simulation.

Before the participants proceeded with the task, they were given a brief instruction explaining the aim and the structure of the task. The participants then joined their group and proceeded with the task. Their performance on the computer screen was recorded. Participants could leave the study anytime they wish. Personal identifiers in the data were anonymized and actual participant names were replaced with pseudo names when transferring printed self-report forms to digital files. Participants were offered a free lunch for participation.

## 4.5 Measures

### 4.5.1 The Metacognitive Awareness Inventory questionnaire

In this study, metacognition was measured by the Metacognitive Awareness Inventory (MAI). The MAI was developed by Schraw & Dennison (1994) and is one of the most commonly used self-reports to measure Metacognition (Norman et al., 2019). It consists of a 52-item instrument (Appendix 1) that in general measures two factors: Knowledge of Cognition and Regulation of Cognition. Subsumed under those two factors are eight subcomponents of metacognition represented in eight scales that all the MAI items distribute across. The eight scales are: *(1) declarative knowledge (DK), (2) procedural knowledge (PK), (3) conditional knowledge (CK),* *(4) planning (P), (5) information management strategies (IMS), (6) comprehension monitoring (CM), (7) debugging strategies (DS), and (8) evaluation (E).* The rating for each item was 100-mm bi-polar scale ranges from 0 (not true of me) to 100 (very true of me). Participants recorded their responses by drawing a slash across the rating scale at a point that best described how true or false the statement was about them.

MAI was selected in this research because it is one of the three existing adolescence adult measures of psychometric reliability which exclusively focuses on metacognitive awareness. The MAI has been perceived to be practical, efficient and easy to use, which can be delivered in both face to face and online classes (Baker & Cerro, 2000).

In the present study, the MAI scale was found to have adequate internal consistency that supported the reliability of eight sub-scales belonging to metacognitive knowledge and metacognitive regulation. Reliability (Cronbach alpha scores) was calculated for each MAI sub-scale and presented as follows: Declarative Knowledge scale ($\alpha = .776$), Procedural Knowledge scale ($\alpha = .826$), Conditional Knowledge scale ($\alpha = .785$), Planning scale ($\alpha = .856$), Information Management Strategies scale ($\alpha = .835$), Comprehension Monitoring scale ($\alpha = .795$), Debugging Strategies scale ($\alpha = .627$), Evaluation scale ($\alpha = .77$).

### 4.5.2 CPS performance on *Tailorshop* simulation

CPS performance was measured in terms of (a) perceived individual CPS performance, (b) perceived group CPS performance, (c) objective group CPS performance. After finishing the Tailorshop task, participants marked their perceived individual performance by rating their individual performance on the scale from 1 (I performed very poor) to 10 (I performed very high). Similarly, participants also marked their perceived group performance by answering the question "How was your group's performance during the task?" and rate on the scale from 1 (We performed very poor) to 10 (We performed very high).

In order to measure objective group CPS performance, the company value in *Tailorshop* simulation was selected as the key variable. The most preferable approach would be to obtain the company value after every month. However, the company monthly values are not independent because the value in one month depends on the previous month's value. This fact violated the assumption of uncorrelated errors (Danner, Hagemann, Holt, et al., 2011). Hence, as an alternative approach, the trends of the company value was used as a group's performance indicator (Funke, 1983). These trends are binary variables. The Excel built-in functions were utilized to demonstrate them. If the company value between two consecutive months increases, the trend was considered “positive”, computed as (+1), and was assigned a “true” label. If the company value decreases on the other hand, the trend was “zero”, computed as (0) and labelled “false”. Eventually, sum of the monthly trends was calculated as the group’s performance indicator, ranging from 0 to 12.

## 4.6 Data and analysis

### 4.6.1 Data preparation

In order to explore the relationship between MAI and CPS performance, the data of groups participated were obtained. Originally, there were simulation data of 26 groups in the study. However, due to technical problems, simulation data of 1 group was omitted and excluded from the analysis. As a result, data of 25 groups were obtained. Each group included 3 members and was totaled of 75 individuals. The data was collected from 2 main sources: Metacognitive Awareness Inventory (MAI) questionnaires and Tailorshop log data in microworld server.

*Questionnaire data*

Individuals’ response to MAI questionnaires was entered into excel file where the columns show 52 Metacognitive Awareness Inventories distributed across eight scales of metacognitive subcomponents. The lines in the excel file show participants identity (ID and alias), group number they belong to, and scores for each and every MAI item. Subsequently, the data was imported into SPSS 20 for further analysis.

*Log data*

The log data represents the group objective CPS performance scores. Data was extracted from the Microworld server which was stored after the participants finished the task. It was later entered into an excel file and imported to SPSS 20 for further analysis.

### 4.6.2 Data analysis

Prior to the statistical analysis, scores for each sub-component of metacognition was calculated. The MAI self-report includes 52 items distributed across 8 sub-components of metacognition, in which the average of all items belonging to a specific component was taken. After that, the distribution of the variables was checked. The variables include CPS performances on different levels and average score of each and every 8 metacognition sub-components. As a result, all variables in the study displayed normal distribution: Perceived Individual CPS Performance (Skewness:-0.932; Kurtosis:0.199), Perceived Group CPS Performance (Skewness:-1.02; Kurtosis:0.42), Objective Group CPS Performance (Skewness:0.11; Kurtosis:-1.5), Average Declarative Knowledge (Skewness:-0.33; Kurtosis:-0.43), Average Procedural Knowledge (Skewness:-0.85; Kurtosis:1.4), Average Conditional Knowledge (Skewness:-0.83; Kurtosis:0.65), Average Planning (Skewness:-1.14; Kurtosis:1.35), Average Information Management Strategies (Skewness:-1.13; Kurtosis:1.6), Average Comprehension Monitoring (Skewness:-0.97; Kurtosis:0.5), Average Debugging Strategies (Skewness:-0.84; Kurtosis:0.82), Average Evaluation (Skewness:-0.5; Kurtosis:-0.77).

*Data analysis for RQ1: GEE*

Data for the first research question was analyzed using a generalized estimating equation (GEE) with a robust covariance estimator, which models the independent variables (e.g., Average Declarative Knowledge, Average Comprehension Monitoring) as predictors of dependent variable in this study, individual perceived performance. GEE is selected in this study because it provides a general method for analyzing clustered variables and does not imply strict assumptions of traditional regression models (Diggle et al., 1995; Liang & Zeger, 1986; Zeger & Liang, 1986). Group numbers are specified as clustered variables that each participant belong to. After cleaning the data, there are 25 clusters in total. Additionally, intracluster dependency correlation matrix was also specified as one of GEEs' requirements (Liang & Zeger, 1986; Zeger & Liang, 1986). In this research, *exchangeable correlation matrix* was specified which represents constant intracluster dependency that off-diagonal elements of the correlation matrix are equal (Collins et al., 2017; Daniels et al., 2016; Morales et al., 2017). Because individual perceived performance, the dependent variable, was treated as continuous, the researcher conducted testing for *normal*, *gamma*, and *inverse Gaussian* distributions with *logarithmic* and *identity link* functions to select the best-fitting models (Garson, 2012).

*Data analysis for RQ2 and RQ3: Best Linear Unbiased Predictors function*

The research's data for RQ2 and RQ3 consists of variables measured at different levels. It comprises of individual students nested within groups. In order to describe the interrelationships among them, multilevel models were employed (Croon & Van Veldhoven, 2007; Hofmann, 1997). There are two common situations in multilevel modelling, macro-micro and micro-macro situations (Croon & Van Veldhoven, 2007; Snijders & Bosker, 1999). The current research's situation fell to a micro-macro multilevel type. A micro-macro situation is determined by a dependent variable defined at the higher group level (group objective CPS performance in this research), is predicted on the basis of independent variables measured at the lower individual level (individual MAI scores and perceived group performance). According to Croon & Van Veldhoven (2007), one of the simplest ways to obtain good estimates of the regression parameters in this situation would be aggregating the individual averaged score to the group level. However, this analysis will yield biased estimates of the regression parameters. Therefore, the Best Linear Unbiased Predictors (BLUP) function was utilized in order to yield unbiased estimates. BLUP function regresses the scores on the groups' unbiased means via adjusted predictors. After the unbiased means are estimated, they can be used in a correlation analysis with other group-level variables (Croon & Van Veldhoven, 2007; Lu et al., 2017). In this research, the researcher implemented BLUP function by using MicroMacro Multilevel package from Lu et al. (2017) in R Version 3.3.3 (R Core Team, 2008). After BLUP scores were calculated, Pearson correlation coefficient was run to measure linear correlation between group-level MAI scores or group perceived CPS performance and group objective CPS performance.

# 5 Results

The present study focuses on examining the relationship between facets of metacognition and perceived individual performance (RQ1) and objective group performance (RQ2) in collaborative CPS. It also investigates the accuracy of calibration of performance in collaborative CPS by observing the relationship between group perceived performance and group objective performance (RQ3). This section presents the results of the data analysis according to the order of those three research questions.

## *5.1 RQ1: Is there any relationship between facets of metacognition and perceived individual performance in collaborative complex problem solving?*

Table 1 reports the quasi-likelihood under the independence criterion (QIC) testing results yielded from normal, gamma, and inverse Gaussian distributions with logarithmic and "identity (linear) link" functions. The researcher reports results from the GEE model using an inverse Gaussian distribution with a logarithmic function (Table 2) because it yielded the lowest QIC value, meaning it was the best fitting.

**Table 1.**

***QIC results yielded from different combinations of distribution and link function using Exchangeable correlation structure in GEE analysis***

| Combination of Distribution – Link function | Quasi Likelihood under Independence Model Criterion (QIC)b | Corrected Quasi Likelihood under Independence Model Criterion (QICC)b |
|---|---|---|
| ***Normal – Identity*** | 389.964 | 393.578 |
| ***Normal – Log*** | 389.589 | 394.053 |
| ***Gamma – Identity*** | 28.278 | 30.713 |
| ***Gamma – Log*** | 25.710 | 32.285 |
| ***Inverse Gaussian – Identity*** | 155.108 | 79.053 |
| ***Inverse Gaussian – Log*** | 17.978 | 21.566 |

**Table 2.**

***GEE results model using an Inverse Gaussian Distribution with a Log Function predicting Perceived Individual Performance (n = 75)***

| Metacognitive Awareness Inventory | Coefficient | SE | Sig. | 95% Wald Confidence Interval | |
|---|---|---|---|---|---|
| | | | | Lower | Upper |
| ***Declarative Knowledge (DK)*** | 0.002 | 0.0071 | 0.765 | -0.012 | 0.016 |
| ***Procedural Knowledge (PK)*** | 0.020 | 0.0108 | 0.058 | -0.001 | 0.042 |
| ***Conditional Knowledge (CK)*** | -0.012 | 0.0110 | 0.263 | -0.034 | 0.009 |
| ***Planning (PL)*** | 0.004 | 0.0070 | 0.592 | -0.010 | 0.017 |
| ***Information Management Strategies (IMS)*** | -0.007 | 0.0095 | 0.467 | -0.026 | 0.012 |
| ***Comprehension Monitoring (CM)*** | 0.009 | 0.0092 | 0.334 | -0.009 | 0.027 |
| ***Debugging Strategies (DS)*** | 0.026* | 0.0131 | 0.044 | 0.001 | 0.052 |
| ***Evaluation (EV)*** | -0.027** | 0.0074 | 0.000 | -0.042 | -0.012 |

From the results reported in Table 2, there are two metacognitive regulation components: debugging strategies component and evaluation component exhibiting statistically significant association with perceived individual performance in collaborative CPS. In which awareness of debugging strategy shows positive relationship ($\beta = 0.026$, $p = 0.044$) and awareness of evaluation shows negative relationship ($\beta = -0.027$, $p = 0.000$) with the perception of individuals on their performance. No correlation was observed between three subcomponents of metacognitive knowledge (declarative, procedural, conditional) and perceived individual performance. Likewise, the rest 3 metacognitive regulation components: planning, information management strategies, comprehension monitoring showed no statistically significant predictors in the model.

## 5.2 *RQ2: Is there any relationship between facets of metacognition and objective group performance in collaborative complex problem solving?*

For the second research question, to specify the dimensions in which metacognition is related to actual group CPS performance, BLUP function was used to regress the MAI scores into the groups' unbiased means via adjusted predictors. After BLUP scores were calculated, Pearson correlation coefficient was run to measure linear correlation between group-level MAI scores, for each of the various components of metacognition, and group objective CPS performance. The results are shown in Table (3).

**Table 3.**

***Correlation Coefficients for group level MAI scores and Objective Group Performance (n = 25)***

| Metacognitive Awareness Inventory | Objective Group Performance in Collaborative CPS | |
|---|---|---|
| | Pearson Correlation | Sig. (2-tailed) |
| ***Declarative Knowledge (DK)*** | -0.065 | 0.761 |
| ***Procedural Knowledge (PK)*** | -0.179 | 0.402 |
| ***Conditional Knowledge (CK)*** | -0.403 | 0.051 |
| ***Planning (PL)*** | -0.133 | 0.535 |
| ***Information Management Strategies (IMS)*** | -0.303 | 0.149 |
| ***Comprehension Monitoring (CM)*** | -0.448* | 0.028 |
| ***Debugging Strategies (DS)*** | -0.470* | 0.020 |
| ***Evaluation (EV)*** | -0.271 | 0.200 |

The results in Table (3) show that all the facets of metacognition had negative correlation with collaborative CPS performance. Negative and significant correlations were found between collaborative CPS performance and two regulatory facets of metacognition: Comprehensive Monitoring ($r = -0.45$, $p = 0.028$) and Debugging Strategies ($r = -0.47$, $p = 0.020$). The strength of these two correlations are medium (-0.3 to -0.5). The researcher did not find an association between awareness of metacognitive knowledge components and actual performance of groups.

No correlation was observed between the rest 3 metacognitive regulation subcomponents (planning, information management strategies, evaluation) and objective group performance.

### *5.3 RQ3: How accurate is student's calibration of performance in collaborative complex problem solving? The relationship between perceived group performance and objective group performance*

In order to examine the accuracy of calibration of performance in collaborative complex problem-solving, the researcher calculated the Pearson correlation coefficients for perceived group performance scores and objective group performance scores. Perceived group performance scores were calculated from BLUP function to result regressed scores from individual level to group level's unbiased estimates. The linear correlation results for the two performances are presented in Table (4).

**Table 4.**

***Correlation coefficients for Perceived Group Performance (after running BLUP) and Objective Group Performance (n = 24)***

| | **Objective Group Performance in Collaborative CPS** | |
|---|---|---|
| | **Pearson Correlation** | **Sig. (2-tailed)** |
| ***Perceived Group Performance in collaborative CPS (after running BLUP)*** | 0.208 | 0.331 |

The results showed no correlation between perceived group performance and objective group performance in collaborative CPS.

# 6 Discussions

The study examines the relationships between facets of metacognition and collaborative CPS performance concerning with both perceived individual and objective group performances. Overall, the results showed that debugging strategies are indicator of increased individual perception of performance whereas evaluation plays an influential role in decreasing individual perceived performance. Objective group performance was negatively influenced by metacognitive comprehensive monitoring and debugging strategies. No correlation was found between perceived group performance and objective group performance. This section presents and discusses these results in the order of the 3 research questions.

## *6.1 RQ1: Is there any relationship between facets of metacognition and perceived individual performance in collaborative complex problem solving?*

Results of the analysis showed that debugging strategies are indicator of increased individual perception of performance whereas evaluation plays an influential role in decreasing perceived individual performance. Firstly, regarding the relationship between debugging strategies and perceived individual performance, it can be interpreted that individuals who believed to possess better debugging strategies skill might have performed better in collaborative complex problem-solving conditions as reflected in perceived performance scores. The findings of the present study are consistent with previous research suggesting that ratings of self-report could be inflated or deflated as a result of self-perceived or actual ability (Hoffman & Schraw, 2010; Solomon, 1984; Walczyk & Hall, 1989). The influence of metacognitive regulatory strategies on performance perception can be interpreted from the self-esteem point of view (Ots, 2013). According to Rosenberg et al. (1995), the perception of one's own performance can be referred as his or her attitude toward him or herself, influenced by affective and rational aspects which in this case, metacognitive facets. Therefore, one way to explain the results can be that the subjects in the present study were well aware of their metacognition, debugging strategies in particular, and that awareness helped them reflect their actual potential. In contrast, another explanation could be that problem-solvers were less able to understand their metacognitive debugging strategies and subsequently affected the low perception of their potential.

Secondly, regarding the relationship between evaluation and perceived individual performance, it can be interpreted that individuals who perceived that they were competent in their evaluation skill might have performed worse, or vice versa, in collaborative complex problem-solving conditions as reflected in perceived performance scores. This is again in line with prior researches regarding self-perceived ability could influence subjects' ratings of self-report (Hoffman & Schraw, 2010; Solomon, 1984; Walczyk & Hall, 1989). One way to explain this is that subjects' perception of performance exhibited underestimation by positioning themselves below their actual attainment level despite being influenced by their awareness of evaluation skill. This interpreted result is important because according to Oleson et al. (2000), self-doubt and underestimation, where individuals inadequately have low expectations toward their performance have been proposed as sources for learning motivation and may help good performers enhance their efforts to succeed the task at hand. Another way to explain this finding is that participants showed incompetence in their evaluation skill by marking it high before the task but contradictorily providing over-estimations on their performance after the task. While the tendency to underestimate oneself may support one's task performance by enhancing motivation, unawareness of one's own incompetence together with overestimations may weaken problem-solvers potential to improve their performance (Ots, 2013).

Regarding collaborative context of the present study, Gutierrez & Price (2017) suggested that subjects tend to attribute their performance judgements to social interaction, which indicates that by interacting and comparing the knowledge of their own with their peers, subjects may be influenced to overestimate their performance. This suggestion runs contrary to previous researches, which suggested that group activities, dependent on the quality of social interactions within the group and the goals held by all group members, could actually help improve self-regulation of learning (Hadwin & Webster, 2013). Collaborative group work is a dynamic process where individuals regulate their own performance and their group members' performance through productive interactions (Hadwin et al., 2017). Productive interactions involve various cognitive, emotional, motivational and metacognitive processes aiming to negotiate and align task representations and goals (Järvelä, Järvenoja, et al., 2016). Therefore, these contradictory findings can be explained further that besides metacognitive skills, the interactions among group members including group-level cognitive, motivational and emotional processes might have had even stronger impact on individuals' perceived performance.

Additionally, a meta-analysis of Dent & Koenka (2016) showed that despite the large amount of studies investigating the influence of metacognition and performance, the association seems to be relatively weak (mean, weighted correlation of r=.20), in which it was stronger when online self-report was employed comparing to the offline one. In the present study, MAI was employed to measure metacognition using offline self-report. Therefore, this might be another reason to explain the contradictory results and non-significant findings of other metacognition knowledge and regulatory facets in relation with perceived individual performance in this study.

Overall, this section reports the results for research question 1 investigating relationship between facets of metacognition and perceived individual performance in collaborative CPS. The results have been discussed by consolidating the views related to individuals' perception of their performance influenced by their perception of themselves and their actual ability in collaborative problem-solving context.

## *6.2 RQ2: Is there any relationship between facets of metacognition and objective group performance in collaborative complex problem solving?*

Results of the second research question showed significantly negative correlation between awareness of metacognitive comprehensive monitoring, debugging strategies and objective group performance in collaborative CPS. This result is consistent with the study of Toy (2007) who found that regulation of cognition measured by MAI was negatively correlated with subjects' scores on complex, ill-structured problem-solving tasks. While metacognitive awareness is hypothesized to enhance performance in solving problems that are complex and ill-structured (Flavell, 1979; Jacobs & Paris, 1987), the present result contradicted with several existing studies (e.g., Shin et al., 2003) which showed that metacognitive regulation had positive correlation with complex problem-solving performance. Complex problem-solving demands both cognitive effort and metacognitive skills such as planning different strategies, monitoring activities to regulate cognition effectively (Jonassen, 1997). However, empirical studies have not reached a consensus on the role of metacognition and performance in general (Toy, 2007). On the one hand, several studies suggest that metacognition does strongly influence learning and problem-solving performance (Goos et al., 2002; Hurme et al., 2006; Schraw & Dennison, 1994a). On the other hand, other studies reported that metacognition was not a predictor of performance (Bendixen & Hartley, 2003; Pintrich et al., 1991) or even had negative correlation (Sperling et al., 2004). This is in line with a meta-analysis of Dent & Koenka (2016) showed that despite the large amount of studies investigating the influence of metacognition and performance, the association seems to be relatively weak (mean, weighted correlation of $r=.20$).

Furthermore, the negative correlation between metacognitive monitoring comprehension or debugging strategies and problem-solving performance could be explained in terms of participants' overconfidence in their performance. Studies have reported that subjects, regardless of ages, are commonly overconfident in their metacognitive skills when performing a task (Glenberg & Epstein, 1987; Pressley & Ghatala, 1990). For instance, Glenberg & Epstein (1987) studied reader's self-assessment of their comprehension monitoring skill for reading tasks. In consistency with their previous studies, they found that readers' beliefs of their monitoring skill are repeatedly off the mark in relation with their reading test performance. The conclusion was drawn that readers have very poor comprehension monitoring skill. Additionally, study of Lin et. al (2002) investigating adults' comprehension and task performance had found that older adults showed problems in their debugging strategies such as rereading. The older adults were found to be as able to detect the problem but they did not purposefully make effort to reread the problems. In that sense, this present study's result is important because it helps identify participants' major obstacles in solving problems. That hypothetically, during the task, problem solvers may believe that they have achieved task's comprehension and stop processing further the text, or stop trying different strategies to fix their on-going problems, which eventually affected their final performance (Glenberg & Epstein, 1987; Lin et al., 2002).

Furthermore, the present study was different from the others when considering its context. The correlation between metacognitive awareness and problem-solving performance was investigated not individually but in collaborative work, which required participants to work together in a group of three, share their understanding and effort to come to a solution. The results of the present study are in line with the study of Pao (2014) who also found that subjects working collaboratively in peers were surprisingly found to have greater overconfidence comparing to individual condition group. This could be interpreted that peer collaboration did not appear to improve metacognitive monitoring but instead became counterproductive in inflating participants' confidence. This result in a way helps hypothesizing the finding of the current study that collaborative aspect didn't have much influence on the relationship between metacognitive regulation and performance and may even had an adverse impact.

On the other hand, a study of Dindar et al. (2020) similarly investigated the influential role of metacognitive experiences, measured during the task, on perceived and objective task performance in collaborative problem solving situation. Significant relationships were found between various metacognitive experiences and perceived and objective performance, both on individual and group levels. These findings might help explain the insignificant findings of the present study, in which no significant relationships were found between metacognitive knowledge, or information management strategies, planning, evaluation and objective group performance.

Metacognition that is measured as a trait, reported by individuals of their metacognitive awareness before doing the tasks (Coutinho et al., 2005), might not be a predictive of performance compared to measuring facets of metacognition during task. Thus, it seems that metacognition can be considered as dynamic process rather than only a static aptitude. Additionally, MAI measuring metacognition in offline self-report form in this study, found to result weaker correlation between metacognition and performance comparing to online self-report (Dent & Koenka, 2016) could also be another reason for non-significant findings in this study.

In terms of complexity, the problem in this study is regarded as a complex problem which involves a numerous variables having mutual dependencies and connectivity (Funke, 2010). A further hypothesis regarding the negative correlation between the subjects' belief in their regulation of cognition skills and objective problem solving performance may be because the problems were too difficult and complex, that despite the belief and confidence in metacognitive regulation skills, due to task complexity, subjects might have not invested further effort in understanding or employing strategies to solve the problems, which then reflected low scores on their actual performance (Coutinho, 2006).

In conclusion, the findings of the second research question of this study addresses and contributes to the overall picture regarding the contradictory findings in the influence of metacognition measured subjectively from the participants' beliefs and objective performance in general or in the context of collaborative complex problem-solving. Possibility of metacognition measured as a trait or as a dynamic process was also discussed. Furthermore, it suggests that the influence of metacognition measured subjectively in relation with actual performance in general is still far more to be understood and call for more studies regarding the influence of metacognition specifically on collaborative complex problem-solving performance.

### *6.3 RQ3: How accurate is student's calibration of performance in collaborative complex problem solving? The relationship between perceived group performance and objective group performance.*

Regarding the results of the third question, no correlation was observed between perceived group performance and objective group performance. One possible explanation of the result could be that subjects may have not considered the scale measuring their perceived performance as a factor determining their performance judgements; which could be due to the complexity of the task that subjects did not have clear awareness of the measurement standard of their performance.

As having been defined, calibration is the process of matching perception of performance with actual level of performance (Nietfeld et al., 2005). It is often referred as an arrival measure of absolute metacognitive monitoring accuracy when confidence judgements are compared with the actual performance (Nietfeld et al., 2005; Schraw, 1995). Calibration is generally categorized into either calibration of comprehension, referring to subjects' confidence judgement in knowledge acquisition or performing ability before a task, or calibration of performance, referring to subjects' confidence judgement in produced outcome after completing the task (García et al., 2016; Nietfeld et al., 2005). Studies have showed calibration of performance tends to be more accurate than calibration of comprehension (Glenberg & Epstein, 1985, 1987; Maki et al., 1990). The third research question in this study addresses calibration of performance.

Calibration of performance plays an important role in informing individuals about their task situation and induces their further effort in controlling and using regulatory strategies to enhance their problem-solving performance (Alexander, 2013; Everson & Tobias, 2001). In general, the difference between objective performance and self-reports perceived performance have been studied within the calibration of performance literature. The majority of them have reported negative relationship between perceived performance and objective performance, meaning that subjects have tendency to be overconfident when rating their own performance (Burson et al., 2006; Hoffman & Schraw, 2010; Schraw et al., 1993).

Regarding problem-solving task, a study of García et al. (2016) also investigated post-performance calibration accuracy in students' problem solving and found that individuals showed little calibration and overconfidence. Moreover, their calibration of performance was reported to be stable across different type of problems, in which the level of judgments and actual performance remain the same. Similarly, previous studies in problem-solving, mathematics problem solving, conducted for students and adults have also reported the same results (Bol et al., 2005, 2010; Hacker et al., 2008; Özsoy, 2012).

Interestingly, a study of Dindar et al. (2020) investigating the influential role of metacognitive experiences on perceived and objective task performance in collaborative problem solving situation, observed a positive relationship between groups' collective judgment of confidence and objective collaborative problem-solving performance, meaning that groups were accurate in their calibration of performance. While it is common for subjects to make inaccurate task performance judgement due to misinterpretation of cues acquired during task monitoring (Koriat, 2015; Ackerman & Goldsmith, 2011), the study of Dindar et al. (2020) has found that the CPS simulations (e.g., Tailorshop) might have done a good job in providing cues in the form of online feedback that helped participants make accurate performance judgement (Ariel, Dunlosky, & Bailey, 2009; Koriat, 1997), and as a result increase calibration of performance accuracy.

Regarding collaborative aspect in this research, a study of (Kirschner et al., 2011) indicated that when a team comprises of members collaboratively working on a task, depending on the type of the task and its complexity, cognitive demands and group performance will increase or decrease accordingly. Therefore, assessing complex problem performance in a collaborative context may have been cognitively demanding that participants found it overwhelming, and as a result, accuracy of calibration of performance may have been diminished (Bol et al., 2012).

Another reason to explain no significant result of the present study could be that participants' performance judgements regardless of skill levels are subject to similar degrees of error. In this sense, although studies have argued that people tend to be overconfident in the task performance, and better metacognitive awareness performers calibrate more accurately (Hoffman & Schraw, 2010; Schraw et al., 1993), this however, is not always the case but depends also on the nature of the task (Burson et al., 2006). In a series of comparison study, Burson and colleagues (2006) showed that on moderately difficult tasks, higher performing or lower performing participants, categorized based on metacognitive awareness and performance, show very little calibration accuracy difference. However, on more difficult tasks, higher performers are interestingly less accurate than lower performers in their judgments of performance. Hence, the present study suggests a need for more empirical investigations regarding calibration of performance accuracy particularly in the context of complex problem-solving task.

In all, this section presents the results of the third research question: the accuracy of participants' calibration of performance in collaborative CPS, determined by the difference between perceived group performance and objective group performance. The results showed no correlation between the two type of performances. The section reviews previous studies relevant to the calibration paradigm and addresses that accuracy and stability of calibration of performance could also depend on other variables particularly task complexity, CPS task's cues, and collaborative aspect of the group task.

# 7 Evaluation

This section evaluates the current study by discussing validity, reliability, and ethical issues in quantitative research method.

## 7.1 Validity

Validity is important because it ensures a research's effectiveness by accurately addressing and representing the features that it aimed to explain or measuring what it aimed to measure (Cohen et al., 2018). Validity is enhanced by staying true to the assumptions corroborating the statistics handling in the research, careful sampling, and suitable research instrumentation. Additionally, different types of validity need to be addressed, including construct validity, statistical conclusion validity of the measures used, and avoiding a range of risks to internal and external validity (Winter, 2000).

*Construct validity.* Construct validity refers to an extent a measure or instrument for data collection to be consistent with the theoretical context of the research (Cohen et al., 2018). In the present study, the theories of metacognitive awareness and CPS have been grounded by a strong body of research in metacognition originated by Flavell (1979) and a rich amount of studies in CPS proposed by (Dörner & Funke, 2017). Instruments for metacognitive related data collection was utilized by the MAI scale which was proved to be one of the most common instruments for adults' measurement of metacognition (Schraw & Dennison, 1994a). For complex problem-solving, Tailorshop simulation is also a common task being employed to measure objective CPS performance (Fischer et al., 2017). Therefore, the instruments and theoretical context are well grounded and consistent for a valid study construct in investigating the relationship between participants metacognitive facets and collaborative complex problem-solving performance.

*Statistical conclusion.* Statistical conclusion validity aims to demonstrate valid statistical procedures and calculation in quantitative research (Cohen et al., 2018). In the present study, statistical procedures were followed by checking normal distribution of the variables of metacognitive facets. Additionally, variables were measured at different levels with individual scores nested in groups. As a result, multilevel modelling and BLUP function were employed to regress group's unbiased means via adjusted predictors. These statistical analysis protocols were utilized and implemented in accordance with the data and the context of the research to ensure its validity.

*Internal validity.* Internal validity aims to demonstrate that the explanation of a research phenomenon can actually be sustained by the data and the findings must describe accurately the phenomena being researched (Cohen et al., 2018). There are several components concerning internal validity in this research. Firstly, instrumentation reliability, the MAI scale selected in this research was proved to have psychometric reliability (Baker & Cerro, 2000). The MAI scale was introduced in prior to the task which left participants time to focus and maintain a high level of concentration. Tailorshop microworld instrument used for measuring objective performance, was considered to be non-intrusive and have been utilised by a sound body of research (Danner, Hagemann, Holt, et al., 2011). Secondly, regarding history, in order to minimize other events apart from the main research intervention during the study, participants were gathered in groups and in quiet and well-facilitated room with least distractions as possible, allowing them to focus on the task. Thirdly, selection bias was also minimized by recruiting participants from various backgrounds with volunteer base and grouping was randomized.

*External validity*. External validity aims to demonstrate the effects to which populations or settings can be generalized (Cohen et al., 2018). There are several components concerning external validity in this research. Firstly, regarding explicit description of independent variables, independent variables were considered to be adequately addressed by the researcher, in which MAI inventories are independent variables measuring participants awareness of their metacognition. Secondly, regarding operationalization of dependent variables, objective complex problem-solving performance assessed by Tailorshop may be generalized in daily CPS but not fully in a real life working complex problem-solving context. Thirdly, regarding representativeness of target populations, while participants can represent university students in collaborative complex problem-solving, they may not be generalized to represent the adult population. However, regarding ecological validity, participants' behaviour observed in this context could be generalized to a larger context of adults behaviour in collaborative CPS situations.

## 7.2 Reliability

Research's reliability refers to its dependability, consistency and replicability over time. It aims to demonstrate that if the research were to be conducted on a similar group of respondents with similar instruments and context, then similar results would be obtained (Cohen et al., 2018). On principal, there are three types of reliability: stability, equivalence and internal consistency. In the present research, the test was conducted one time, and there was no re-test after the first test had been finished. Therefore, the risk of errors in testing over-time could be minimized.

Regarding the research's stability, statistical significance of the correlation coefficient was found using Pearson correlation method in the present research. Two out of three research questions reported significant correlation with a standard level of significance $p < 0.05$, for both correlational and multi-level. Therefore, reliability is to be guaranteed in this aspect.

Regarding reliability as internal consistency, Cronbach alpha is used as a measure to provide coefficient of inter-item correlations, in which the correlation of an item with the sum of all other relevant items in a multi-item scale. In the present study, the Metacognitive Awareness Inventory (MAI) scale display good internal consistency based on Cronbach alpha scores in eight Metacognitive dimension scales: Declarative Knowledge scale ($\alpha = .776$), Procedural Knowledge scale ($\alpha = .826$), Conditional Knowledge scale ($\alpha = .785$), Planning scale ($\alpha = .856$), Information Management Strategies scale ($\alpha = .835$), Comprehension Monitoring scale ($\alpha = .795$), Debugging Strategies scale ($\alpha = .627$), Evaluation scale ($\alpha = .77$).

## 7.3 Ethical issues

Ethical issues concern principles and practices researchers' conducts' related to the rights and values of the subjects participating in the research (Cohen et al., 2018). Ethical issues may stem from the research topics and methods used to obtain valid and reliable data. Therefore, ethical issues are required throughout the stages of the research.

In the present study, ethical issues had been addressed throughout different stages, from research purpose, procedure to research content. This thesis is a part of the research project of learning, education and technology programme at the university in advancing the understanding of complex problem-solving phenomena. Metacognition is also one of the key study subjects in the programme. Therefore, the thesis inherits the research-focused goals of the research department in understanding learning and problem solving.

Regarding research procedure, respondents' rights and values had been practised in the research. Participants recruitment was voluntarily based. Consents form stating the research aims, protection of participants' rights and data was informed and distributed to the participants at the beginning of each session. Respondents were asked to read, fully comprehend and sign on the form if they agree. Respondents were informed to ask for any clarifications if needed. They could disagree and leave the research at any time if needed.

Regarding data collection, analysis and reporting, participants' records and scores were kept in confidential, both with the digital data from Tailorshop objective performance as well as non-digital MAI scores and perceived performance scores. Additionally, subjects' names in the record were translated into alias in order to avoid bias and secure anonymity of participants.

# 8 Conclusion

Prior researches have showed that metacognition plays an influential role in both perceived task performance (Hoffman & Schraw, 2010; Solomon, 1984; Walczyk & Hall, 1989) and objective task performance (Flavell, 1979; Goos et al., 2002; Hurme et al., 2006; Jacobs & Paris, 1987; Shin et al., 2003).

In this study, metacognition was investigated in detail on the facet level, in which it was operationalised into Metacognitive Awareness Inventories presenting eight subcomponents of metacognitive knowledge and metacognitive regulation (Schraw & Dennison, 1994a). Similarly, task performance was examined more specifically in terms of collaborative complex problem-solving task which not only involves individual complex problem-solving process but also collaborative social aspect (Dillenbourg, 1999; Roschelle & Teasley, 1995).

The results of the present study indicate significant relationships between metacognitive regulation components and perceived individual performance as well as objective group performance in collaborative complex-problem solving context. Therefore, the study's findings reinforce metacognition influential role on task performance particularly in collaborative complex problem-solving context.

Regarding perception on individual CPS task performance, the results indicate positive relationship between debugging strategies and perceived individual performance, but negative relationship between evaluation and perceived individual performance. This supports previous findings that individuals' underestimation helps enhance their effort to succeed the task while overestimation weakens their potential to improve the performance (Oleson et al., 2000).

Regarding objective group performance, the results suggest that students' prior beliefs in metacognitive regulation plays a negatively influential role on objective group performance which reinforces the common findings that participants tend to be overconfident in their metacognitive skills in problem solving (Glenberg & Epstein, 1987; Pressley & Ghatala, 1990). While some studies reported that metacognition has negative correlation (e.g., Sperling et al., 2004), other researches suggest that metacognition does strongly influence problem-solving performance (Goos et al., 2002; Hurme et al., 2006; Schraw & Dennison, 1994a) or even not a predictor of performance at all (Bendixen & Hartley, 2003; Pintrich et al., 1991). Hence, the current study has contributed to overall understanding of whether metacognition plays an influential role in task performance in general and suggests that this relationship is still far more to be understood.

Results of the third research question showed that groups were not calibrated in their performance. This was likely a novel task for many of the students and therefore they might have not had standards to compare with their performance. Additionally, various aspects in group work combined with the complexity of the task are considered the reasons affect the stability of calibration of performance as well as the relationship between metacognition and participants' perceived and objective performance (Cooke et al., 2004; Dillenbourg, 1999; Pao, 2014). Therefore, this work may shed a light on the research direction investigating the role of metacognition in collaborative CPS.

In summary, this study has contributed to the need for more empirical research on the impact of metacognition factors in the collaborative complex problem-solving literature. The use of the MAI self-report has served for the investigation of the awareness of metacognition, revealing the beliefs of students on various dimensions of metacognition. The *Tailorshop* simulation has played a role on simulating the task and evaluating objective performance of collaboration CPS performance. The study has found a negative relation between metacognitive regulation dimensions and objective group performance, whereas in the individual level, metacognitive debugging strategy has been found to positively and metacognitive evaluation has been found to negatively affect individual's self-perception of their performance. Therefore, it is proven that there are relationships between metacognition facets and collaborative CPS tasks performance.

## 8.1 Limitations and future research

The current study has several limitations to be addressed. Firstly, the sample size of the study can be considered modest, although the results exhibited good fit values. Given the small number of studies that investigated these variables in the context of collaborative complex problem-solving, it is imperative that a larger sample size could be considered in the future for research replications.

Secondly, in order to increase construct validity, future researches could consider conducting the research of CPS in a context of a course or a task at work. Placing a research in a real context could increase respondents' motivation and engagement (Cohen et al., 2018) that as a result would produce more accurate data from participants' task performance and assessment.

Thirdly, previous studies having consistent results with research question 1 and research question 3 of this thesis have controlled group forming with members who are considered high performer or low performer measured by their metacognitive awareness scores. This comparative approach would help to explain better the results especially between different metacognitive components and task perceived performance in interpretation of under or overestimation of task performance, as well as broaden understanding of calibration of performance influenced by different variables such as task complexity and collaborative aspect (Burson et al., 2006; Ots, 2013).

## References


Ackerman, P. L., & Wolman, S. D. (2007). Determinants and validity of self-estimates of abilities and selfconcept measures. *Journal of Experimental Psychology: Applied*, *13*, 57–78.

Alexander, P. A. (2013). Calibration: What is it and why it matters? An introduction to the special issue on calibrating calibration. *Learning and Instruction*, *24*, 1–3. https://doi.org/10.1016/j.learninstruc.2012.10.003.

Aljaberi, N. M., & Gheith, E. (2015). University Students' level of metacognitive thinking and their ability to solve problems. *American International Journal of Contemporary Research*, *5*(3), 121–134.

Amelung, D., & Funke, J. (2013). Dealing with the uncertainties of climate engineering: Warnings from a psychological complex problem solving perspective. *Technology in Society*, *35*(1), 32–40. https://doi.org/10.1016/j.techsoc.2013.03.001

Andreas Fischer,Samuel Greiff,Joachim Funk. (2017). *The history of complex problem solving*.

Artz, A. F., & Armour-Thomas, E. (1992). Development of a cognitive-metacognitive framework for protocol analysis of mathematical problem solving in small groups. *Cognition and Instruction*, *9*(2), 137–175.

Baker. (1989). Metacognition, comprehension monitoring, and the adult reader. *Educ. Psychol. Rev*, *1*, 3–38.

Baker, & Anderson, R. I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. *Read. Res. Q*, *17*, 281–294.

Baker, L., & Cerro, L. C. (2000). *3. Assessing Metacognition in Children and Adults*. 49.

Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *Journal of the Learning Sciences*. https://doi.org/10.1207/S15327809JLS0904_2

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*. https://doi.org/10.1207/S15327809JLS1203_1

Bendixen, L. D., & Hartley, K. (2003). Successful learning with hypermedia: The role of epistemological believes and metacognitive awareness. *Journal of Educational Computing Research*, *28*, 15–30.

Bereiter, C., & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. *Cognition and Instruction*, *2*(2), 131–156.

Bereiter, C., & Scardamalia, M. (1987). *The Psychology of Written Composition*. Erlbaum.

Bereiter, C., & Scardamalia, M. (1993). *Surpassing Ourselves: An Inquiry into the Nature and Implications of Expertise*.

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining Twenty-First Century Skills. In *Assessment and Teaching of 21st Century Skills* (pp. 17–66). Springer Netherlands. https://doi.org/10.1007/978-94-007-2324-5_2

Blohm, P. J. (1987). Effect of lookup aids on mature readers' recall of technical text. *Read. Res. Instruct*, *26*, 77–88.

Bol, L., Hacker, D. J., O'Shea, P., & Allen, D. (2005). The influence of overt practice, achievement level, and explanatory style on calibration accuracy and performance. *Journal of Experimental Education*, *73*, 269–290. https://doi.org/10.3200/JEXE.73.4.269-290.

Bol, L., Hacker, D. J., Walck, C. C., & Nunnery, J. A. (2012). The effects of individual or group guidelines on the calibration accuracy and achievement of high school biology students. *Contemporary Educational Psychology*, *37*(4), 280–287. https://doi.org/10.1016/j.cedpsych.2012.02.004

Bol, L., Riggs, R., Hacker, D. J., & Nunnery, J. (2010). The calibration accuracy of middle school students in math classes. *Journal of Research in Education*, *21*, 81–96.

Bouffard, T., Vezeau, C., Roy, M., & Lengelé, A. (2011). Stability of biases in self-evaluation and relations to well-being among elementary school children. *International Journal of Educational Research*, *50*, 221–229. https://doi.org/10.1016/j.ijer.2011.08.003.

Brand, S., Reimer, T., & Opwis, K. (2003). Effects of metacognitive thinking and knowledge acquisition in dyads on individual problem solving and transfer performance. *Swiss Journal of Psychology*, *62*(4), 251–261. https://doi.org/10.1024/1421-0185.62.4.251

Brown, A. (1987). *Metacogntion, executive control, self-regulation, and other more*.

Burson, K. A., Larrick, R. P., & Klayman, J. (2006). Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons. *Journal of Personality and Social Psychology*, *90*(1), 60.

Campion, M. A., Medsker, G. J., & Higgs, C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Pers. Psychol*, *46*, 823–850. https://doi.org/10.1111/j.1744-6570.1993.tb01571.x

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Individual and Group Decision Making* (pp. 221–246). Lawrence Erlbaum Associates.

Carlson, M. P., & Bloom, I. (2005). The cyclic nature of problem solving: An emergent multidimensional problem-solving framework. *Educational Studies in Mathematics*, *58*(1), 45–75.

Cera, R., Mancini, M., & Antonietti, A. (2013). Relationships between Metacognition, Self-efficacy and Self-regulation in Learning. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, *4*(7), 115–141. https://doi.org/10.7358/ecps-2013-007-cera

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145–182.

Cohen, L., Manion, L., & Morrison, K. (2018). Validity and reliability. In L. Cohen, L. Manion, & K. Morrison (Eds.), *Research Methods in Education* (pp. 245–284). Routledge. https://doi.org/10.4324/9781315456539-7

Collins, T. W., Grineski, S. E., Shenberger, J., Morales, D. X., Morera, O., & Echegoyen, L. (2017). Undergraduate research participation improves educational outcomes. *Journal of College Student Development*.

Cooke, N. J., Salas, E., Kiekel, P. A., & Bell, B. (2004). Advances in measuring team cognition. In E. Salas & M. S. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance*. American Psychological Associa-tion. https://doi.org/10.1037/10690-005

Coutinho, S. A. (2006). *The Relationship between the Need for Cognition, Metacognition, and Intellectual Task Performance*. 3.

Coutinho, S., Wiemer-Hastings, K., Skowronski, J. J., & Britt, M. A. (2005). Metacognition, need for cognition and use of explanations during ongoing learning and problem solving. *Learning and Individual Differences*, *15*(4), 321–337. https://doi.org/10.1016/j.lindif.2005.06.001

Cress, U., & Hesse, F. (2013). *Quantitative methods for studying small groups* (pp. 93–111).

Creswell, J. W. (1994). *Research design: Qualitative and quantitative approach*. Publications.

Croon, M. A., & Van Veldhoven, M. J. P. M. (2007). Predicting group-level outcome variables from variables measured at the individual level: A latent variable multilevel model. *Psychological Methods*. https://doi.org/10.1037/1082-989X.12.1.45

Curseu, P. L. (2006). Need for cognition and rationality in decision-making. *Studia Psycho-Logica*, *48*(2), 141.

Daniels, H., Grineski, S. E., Collins, T. W., Morales, D. X., Morera, O., & Echegoyen, L. (2016). Factors influencing student gains from undergraduate research experiences at a Hispanic-serving institution. *CBE—Life Sciences Education*, *15*(3), 30.

Danner, D., Hagemann, D., Holt, D. V., Hager, M., Schankin, A., Wüstenberg, S., & Funke, J. (2011). Measuring Performance in Dynamic Decision Making. *Journal of Individual Differences*, *32*(4), 225–233. https://doi.org/10.1027/1614-0001/a000055

Danner, Hagemann, D., Schankin, A., Hager, M., & Funke, J. (2011). Beyond IQ: A la-tent state-trait analysis of general intelligence, dynamic decision making, and implic-it learning. *Intelligence*, *39*(5), 323–334. https://doi.org/10.1016/j.intell.2011.06.004

de Bruin, A. B., & van Gog, T. (2012). *Improving self-monitoring and self-regulation: From cognitive psychology to the classroom*.

Delclos, V. R., & Harrington, C. (1991). Effects of strategy monitoring and proactive instruction on children's problem-solving performance. *J. Educ. Psychol*, *83*, 35–42.

Dent, A. L., & Koenka, A. C. (2016). The Relation Between Self-Regulated Learning and Academic Achievement Across Childhood and Adolescence: A Meta-Analysis. *Educational Psychology Review*, *28*(3), 425–474. https://doi.org/10.1007/s10648-015-9320-8

Dierdorff, E. C., & Ellington, J. K. (2012). Members matter in team training: Multilevel and longitudinal relationships between goal orientation, self-regulation, and team outcomes. *Personnel Psychology*, *65*(3), 661–703.

Diggle, P. J., Liang, K. Y., & Zeger, S. L. (1995). Analysis of longitudinal data. *Journal of the Royal Statistical Society. Series A, Statistics in Society*, *158*(2), 339.

Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively? *Educational Research Review*, *3*(2), 101–129. https://doi.org/10.1016/j.edurev.2008.02.003

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp. 1–19). Elsevier PP.

Dindar, Järvelä, S., & Järvenoja, H. (2020). Interplay of metacognitive experiences and performance in collaborative problem solving. *Computers & Education*, *103922*.

Dindar, M. (2018). An empirical study on gender, video game play, academic success and complex problem solving skills. *Computers and Education*. https://doi.org/10.1016/j.compedu.2018.05.018

Dinsmore, D. L., & Parkinson, M. M. (2013). What are confidence judgments made of? Students' explanations for their confidence ratings and what that means for calibration. *Learning and Instruction*, *24*, 4–14. https://doi.org/10.1016/j.learninstruc.2012.06.001.

Donker, A. S., de Boer, H., Kostons, D., Dignath van Ewijk, C. C., & van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, *11*, 1–26. https://doi.org/10.1016/j.edurev.2013.11.002

Dörner. (1975). Wie menschen eine Welt verbessern wollten" (how people wanted to improve the world. *Bild der Wissenschaft*, *12*(2), 48–53.

Dörner. (1986a). Diagnostik der operativen Intelligenz" [assessment of operative intelli-gence. *Diagnostica*, *32*(4), 290–308.

Dörner. (1986b). Diagnostik der operativen Intelligenz" [assessment of operative intelli-gence. *Diagnostica*, *32*(4), 290–308.

Dörner. (1997). *The Logic of Failure: Recognizing and Avoiding Error in Complex Situa-tions*. Basic Books.

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

Dörner, Drewes, & Reither, F. (1975). Über das Problemlösen in sehr komplexen Realitätsbereichen" [on problem solving in very complex domains of reality. In W. Tack (Ed.), *Bericht über den 29. Kongress der DGfPs in Salzburg 1974* (Vol. 1, pp. 339–340). hogrefe.

Dörner, & Funke, J. (2017b). Complex Problem Solving: What It Is and What It Is Not. *Frontiers in Psychology*, *8*, 1153. https://doi.org/10.3389/fpsyg.2017.01153

Dörner, Kreuzig, H. W., Reither, F., & Stäudel, T. (1983). *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Huber.

Driskell, J., Salas, E., & Hughes, S. (2010). Collective orientation and team performance: Development of an individual differences measure. *Hum. Factors*, *52*, 316–328. https://doi.org/10.1177/0018720809359522

Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. SAGE Publications.

Efklides, A. (2006). Metacognition and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review*, *1*(1), 3–14.

Eichmann, B., Goldhammer, F., Greiff, S., Pucite, L., & Naumann, J. (2019a). The role of planning in complex problem solving. *Computers & Education*, *128*, 1–12. https://doi.org/10.1016/j.compedu.2018.08.004

Eichmann, B., Goldhammer, F., Greiff, S., Pucite, L., & Naumann, J. (2019b). The role of planning in complex problem solving. *Computers & Education*, *128*, 1–12. https://doi.org/10.1016/j.compedu.2018.08.004

Ellis, D., & Zimmerman, B. J. (2001). Enhancing self-monitoring during self-regulated learning of speech. In H.J. Hartman (Ed.), *Metacognition in learning and instruction* (pp. 205–228). Kluwer.

Engelbrecht, J., Harding, A., & Potgieter, M. (2005). Undergraduate Students' Performance and Confidence in Procedural and Conceptual Mathematics. *International Journal of Mathematical Education in Science and Technology*, *36*(7), 701–712.

Everson, H. T., & Tobias, S. (2001). The Ability to Estimate Knowledge and Performance in College: A Metacognitive Analysis. In Hope J. Hartman (Ed.), *Metacognition in Learning and Instruction* (Vol. 19, pp. 69–83). Springer Netherlands. https://doi.org/10.1007/978-94-017-2243-8_4

Fischer, A., Greiff, S., & Funke, J. (2017). *The history of complex problem solving* (pp. 107–121). https://doi.org/10.1787/9789264273955-9-en

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

Funke. (1988). Using Simulation to Study Complex Problem Solving. *Simulation & Games*, *19*(3), 277–303. https://doi.org/10.1177/0037550088193003

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

Funke. (2003). *Problemlösendes Denken [Problem Solving and Thinking*. Kohlhammer.

Funke. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, *11*(2), 133–142. https://doi.org/10.1007/s10339-009-0345-0

Funke, Fischer, A., & Holt, D. V. (2018). Competencies for Complexity: Problem Solving in the Twenty-First Century. In E. Care, P. Griffin, & M. Wilson (Eds.), *Assessment and Teaching of 21st Century Skills* (pp. 41–53). Springer International Publishing. https://doi.org/10.1007/978-3-319-65368-6_3

Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. In R. J. Sternberg & P. A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms*. Psychology Press.

Funke, J. (2012). Complex Problem Solving. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 682–685). Springer US. https://doi.org/10.1007/978-1-4419-1428-6_685

Funke, Joachim. (2010). Complex problem solving: A case for complex cognition? *Cognitive Processing*, *11*(2), 133–142. https://doi.org/10.1007/s10339-009-0345-0

García, T., Rodríguez, C., González-Castro, P., González-Pienda, J. A., & Torrance, M. (2016). Elementary students' metacognitive processes and post-performance calibration on mathematical problem-solving tasks. *Metacognition and Learning*, *11*(2), 139–170. https://doi.org/10.1007/s11409-015-9139-1

Garner, R. (1987). *Metacognition and Reading Comprehension*. Ablex Publishing.

Garner, R. (1990). When children and adults do not use learning strategies: Toward a theory of settings. *Rev. Educ. Res*, *60*, 517–529.

Garner, R., & Alexander, P. A. (1989). Metacognition: Answered and unanswered questions. *Educ. Psychol*, *24*, 143–158.

Garson, G. (2012). *Generalized linear models and generalized estimating equations*. Statistical Associates Publishing.

Glaser, R., Chi, M. T. H., MTH, G., R., & Farr, M. J. (1988). *Overview in: Chi*.

Glenberg, A. M., & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*, 702–718.

Glenberg, A. M., & Epstein, W. (1987). Inexpert calibration of comprehension. *Memory & Cognition*, *15*, 84–93.

Goos, M., & Galbraith, P. (1996). Do it this way! Metacognitive strategies in collaborative mathematical problem solving. *Educational Studies in Mathematics*, *30*(3), 229–260. https://doi.org/10.1007/BF00304567

Goos, M., Galbraith, P., & Renshaw, P. (2002). Socially mediated metacognition: Creating collaborative zones of proximal development in small group problem solving. *Educational Studies in Mathematics*, *49*(2), 193–223. https://doi.org/10.1023/A:1016209010120

Graesser, A. C., Cai, Z., Hu, X., Foltz, P., Greiff, S., Kuo, B.-C., & Williamson Shaffer, D. (2017). Assessment of Collaborative Problem Solving. *Journal of Intelligence*, *5*(10), 275–285. https://doi.org/10.3390/jintelligence5020010

Greiff. (2012). *Individualdiagnostik komplexer Problemlösefähigkeit [assessment of Complex ProblemSolving ability*. Waxmann.

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

Greiff, Niepel, C., Scherer, R., & Martin, R. (2016). Understanding students' performance in a computer-based assessment of complex problem solving: An analysis of behav-ioral data from computer-generated log files. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2016.02.095

Greiff, S., Fischer, A., Wüstenberg, S., Sonnleitner, P., Brunner, M., & Martin, R. (2013). A multitrait–multimethod study of assessment instruments for complex problem solving. *Intelligence*, *41*(5), 579–596. https://doi.org/10.1016/j.intell.2013.07.012

Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B. (2013a). Complex problem solving in educational contexts—Something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, *105*(2), 364–379. https://doi.org/10.1037/a0031856

Greiff, Stadler, M., Sonnleitner, P., Wolff, C., & Martin, R. (2015). Sometimes less is more: Comparing the validity of complex problem solving measures. *Intelligence*, *50*, 100–113. https://doi.org/10.1016/j.intell.2015.

Greiff, Wüstenberg, S., & Funke, J. (2012). Dynamic Problem Solving: A New Assessment Perspective. *Applied Psychological Measurement*. https://doi.org/10.1177/0146621612439620

Greiff, Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B. (2013b). Complex problem solving in educational contexts—Something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, *105*(2), 364–379. https://doi.org/10.1037/a0031856

Güss, Burger, M. L., & Dörner, D. (2017). The Role of Motivation in Complex Prob-lem Solving. *Frontiers in Psychology*. https://www.frontiersin.org/article/10.3389/fpsyg.2017.00851

Gutierrez, A. P., & Price, A. F. (2017). Calibration between undergraduate students' prediction of and actual performance: The role of gender and performance attributions. *The Journal of Experimental Education*, *85*(3), 486–500.

Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. *Handbook of Metamemory and Memory*, *429455*.

Hadwin, A. F., Jarvela, S., & Miller, M. (2017). Self-regulation, co-regulation and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 83–106). Routledge.

Hadwin, A. F., & Webster, E. A. (2013). Calibration in goal setting: Examining the nature of judgments of confidence. *Learning and Instruction*, *24*(12), 37–47. https://doi.org/10.1016/j.learninstruc.2012.10.001.

Hattie, J. (2013). Calibration and confidence: Where to next? *Learning and Instruction*, *24*, 62–66. https://doi.org/10.1016/j.learninstruc.2012.05.009.

Helms-Lorenz, M., & Jacobse, A. E. (2008). Metacognitive skills of the gifted from a cross-cultural perspective. *Metacognition: A Recent Review of Research, Theory, and Perspectives*, 3–43.

Hertel, G., Konradt, U., & Orlikowski, B. (2004). Managing distance by interdependence: Goal setting, task interdependence, and team-based rewards in virtual teams. *Euro. J. Work Organ. Psychol*, *13*, 1–28. https://doi.org/10.1080/13594320344000228

Hinsz, V., Tindale, R., & Vollrath, D. (1997). The emerging concept of groups as infor-mation processors. *Psychol. Bull*, *121*, 43–64. https://doi.org/10.1037/0033-2909.121.1.43

Hoffman, B., & Schraw, G. (2010). Conceptions of efficiency: Applications in learning and problem solving. *Educational Psychologist*, *45*(1), 1–14.

Hofmann, D. A. (1997). An Overview of the Logic and Rationale of Hierarchical Linear Models. *Journal of Management*, *23*(6), 723–744. https://doi.org/10.1177/014920639702300602

Hurme, T.-R., Palonen, T., & Järvelä, S. (2006). Metacognition in joint discussions: An analysis of the patterns of interaction and the metacognitive content of the networked discussions in mathematics. *Metacognition and Learning*, *1*(2), 181–200. https://doi.org/10.1007/s11409-006-9792-5

Iiskala, T., Vauras, M., & Lehtinen, E. (2004). Socially-shared metacognition in peer learn-ing? *Hellenic Journal of Psychology*, *1*(2), 147–178.

Iiskala, T., Vauras, M., Lehtinen, E., & Salonen, P. (2011). Socially shared metacognition of dyads of pupils in collaborative mathematical problem-solving processes. *Learning and Instruction*, *21*(3), 379–393.

Isohätälä, J., Järvenoja, H., & Järvelä, S. (2017). Socially shared regulation of learning and participation in social interaction in collaborative learning. *International Journal of Educational Research*, *81*, 11–24.

Jacobs, J. E., & Paris, S. G. (1987). Children's metacognition about reading: Issues in definition, measurement, and instruction. *Educational Psychologist*, *22*(3–4), 255–278.

Järvelä, S., Järvenoja, H., Malmberg, J., & Hadwin, A. F. (2013). Exploring Socially Shared Regulation in the Context of Collaboration. *Journal of Cognitive Education and Psy-Chology*. https://doi.org/10.1891/1945-8959.12.3.267

Järvelä, S., Järvenoja, H., Malmberg, J., Isohäätälä, J., & Sobocinski, M. (2016). How do types of interaction and phases of self-regulated learning set a stage for collaborative engagement? *Learning and Instruction*, *43*, 39–51. https://doi.org/10.1016/j.learninstruc.2016.01.005.

Järvelä, S., Järvenoja, H., & Veermans, M. (2008). Understanding the dynamics of motiva-tion in socially shared learning. *International Journal of Educational Research*, *47*(2), 122–135. https://doi.org/10.1016/j.ijer.2007.11.012

Järvelä, S., Kirschner, P. A., Hadwin, A., Järvenoja, H., Malmberg, J., Miller, M., & Laru, J. (2016). Socially shared regulation of learning in CSCL: Understanding and prompting individual-and group-level shared regulatory activities. *International Journal of Computer-Supported Collaborative Learning*, *11*(3), 263–280.

Järvenoja, H., Järvelä, S., & Malmberg, J. (2015). Understanding regulated learning in situative and contextual frameworks. *Educational Psychologist*, *50*(3), 204–219.

Jonassen, D. H. (1997). Instructional design models for well-structured and ill-structured problem-solving learning outcomes. *Educational Technology: Research and Development*, *45*(1), 65–94.

Justice, E. M., & Weaver-McDougall, R. G. (1989). Adult's knowledge about memory: Awareness and use of memory strategies across tasks. *J. Educ. Psychol*, *81*, 214–219.

Kaufman, N. J., Randlett, A. L., & Price, J. (1985). Awareness of the use of comprehension strategies in good and poor college readers. *Reading Psychology: An International Quarterly*, *6*(1–2), 1–11.

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

King, A. (1991). Effects of training in strategic questioning on children's problem-solving performance. *J. Educ. Psychol*, *83*, 307–317.

Kirschner, F., Paas, F., Kirschner, P. A., & Janssen, J. (2011). Differential effects of problem-solving demands on individual and collaborative learning outcomes. *Learning and Instruction*, *21*, 587–599.

Kletzien, S. B. (1988). A comparison of achieving and nonachieving readers' use of comprehension strategies on different reading levels. *Meeting of the American Educational Research Association*.

Kluwe, R. H. (1982). Cognitive knowledge and executive control: Metacognition. In D. R. Griffin (Ed.), *Animal mind – human mind* (pp. 201–224). Springer-Verlag.

Kröner, S. (2001). *Intelligenzdiagnostik per Computersimulation [Assessing intelligence with computer simulations*. Waxmann Verlag.

Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, *73*(1), 13–22. https://doi.org/10.1093/biomet/73.1.13

Lin, L. M., Zabrucky, K. M., & Moore, D. (2002). Effects of text difficulty and adults' age on relative calibration of comprehension. *The American Journal of Psychology*, *115*, 187–198.

Linnenbrink-Garcia, L., Rogat, T. K., & Koskey, K. L. K. (2011). Affect and engagement during small group instruction. *Contemporary Educational Psychology*. https://doi.org/10.1016/j.cedpsych.2010.09.001

Lorch, R. F., Lorch, E. P., & Klusewitz, M. A. (1993). College students' conditional knowledge about reading. *J. Educ. Psychol*, *85*, 239–252.

Lu, M., Zhou, J., Naylor, C., Kirkpatrick, B. D., Haque, R., Petri, W. A., & Ma, J. Z. (2017). Application of penalized linear regression methods to the selection of environmental enteropathy biomarkers. *Biomarker Research*, *5*(1), 9. https://doi.org/10.1186/s40364-017-0089-4

Lucangeli, D., Coi, G., & Bosco, P. (1997). Metacognitive awareness in good and poor math problem solvers. *Learning Disabilities Research & Practice*, *12*(4), 209–212.

Maki, R. H., Foley, J. M., Kajer, W. K., Thompson, R. C., & Willert, M. G. (1990). Increased processing enhances calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 609–616.

Malmberg, J., Haataja, E., Seppänen, T., & Järvelä, S. (2019). Are we together or not? The temporal interplay of monitoring, physiological arousal and physiological synchrony during a collaborative exam. *International Journal of Computer-Supported Collaborative Learning*, *14*(4), 467–490.

Malmberg, J., Järvelä, S., & Järvenoja, H. (2017). Capturing temporal and sequential patterns of self-, co-, and socially shared regulation in the context of collaborative learning. *Contemporary Educational Psychology*, *49*, 160–174.

Malmberg, J., Järvelä, S., Järvenoja, H., & Panadero, E. (2015). Promoting socially shared regulation of learning in CSCL: Progress of socially shared regulation among high-and low-performing groups. *Computers in Human Behavior*, *52*, 562–572.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Acad. Manag. Rev*, *26*, 356–376. https://doi.org/10.2307/259182

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *J. Appl. Psychol*, *85*, 273–283. https://doi.org/10.1037/0021-9010.85.2.273

Miller, P. H. (1985). Metacognition and Attention. In D. L. Forrest-Pressley, E. G. McKinnon, & W. T. G (Eds.), *Metacognition, Cognition, and Human Performance* (pp. 181–221). Academic Press.

Morales, D. X., Grineski, S. E., & Collins, T. W. (2017). Increasing Research Productivity in Undergraduate Research Experiences: Exploring Predictors of Collaborative Faculty–Student Publications. *CBE—Life Sciences Education*, *16*(3), ar42. https://doi.org/10.1187/cbe.16-11-0326

Nair, K. U., & Ramnarayan, S. (2000a). Individual Differences in Need for Cognition and Complex Problem Solving. *Journal of Research in Personality*, *34*(3), 305–328. https://doi.org/10.1006/jrpe.1999.2274

Nair, K. U., & Ramnarayan, S. (2000b). Individual Differences in Need for Cognition and Complex Problem Solving. *Journal of Research in Personality*, *34*(3), 305–328. https://doi.org/10.1006/jrpe.1999.2274

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, *95*(1), 109.

Neubert, J. C., Kretzschmar, A., Wüstenberg, S., & Greiff, S. (2014). Extending the Assessment of Complex Problem Solving to Finite State Automata. *European Journal of Psychological Assessment*, *31*(3), 181–194. https://doi.org/10.1027/1015-5759/a000224

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2005). Metacognitive monitoring accuracy and student performance in the postsecondary classroom. *Journal of Experimental Education*, *74*(1), 22–7.

Norman, E., Pfuhl, G., Sæle, R. G., Svartdal, F., Låg, T., & Dahl, T. I. (2019). Metacognition in Psychology. *Review of General Psychology*, *23*(4), 403–424. https://doi.org/10.1177/1089268019883821

Norman, Elisabeth, Pfuhl, G., Sæle, R. G., Svartdal, F., Låg, T., & Dahl, T. I. (2019). Metacognition in Psychology. *Review of General Psychology*, *23*(4), 403–424. https://doi.org/10.1177/1089268019883821

O'Donnell, A. M., & Hmelo-Silver, C. E. (2013). Introduction What is Collaborative Learning?: An Overview. In C. E. Hmelo-Silver, C. A. Chinn, C. K. K. Chan, & A. M. O'Donnell (Eds.), *The international handbook of collaborative learning* (pp. 13–28). Routledge.

O.E.C.D. (2014). Assessing Problem-Solving Skills in PISA 2012. In P.I.S.A. (Ed.), *PISA 2012 Results: Creative Problem Solving: Vol. V* (pp. 25–47). OECD Publishing. https://doi.org/10.1787/9789264208070-en

OECD. (2014). Assessing Problem-Solving Skills in PISA 2012. In PISA (Ed.), *PISA 2012 Results: Creative Problem Solving (Volume V)* (pp. 25–47). OECD Publishing. https://doi.org/10.1787/9789264208070-en

Opwis, K. (1996). Reflexionen über eigenes und fremdes Wissen. [Reflections about own and other people's knowledge. In F. Klix & H. Spada (Eds.), *Enzyklopädie der Psychologie* (pp. 269–401). Theorie und Forschung.

Ots, A. (2013). Third graders' performance predictions: Calibration deflections and academic success. *European Journal of Psychology of Education*, *28*(2), 223–237.

Özsoy, G. (2012). Investigation of fifth grade students' mathematical calibration skills. *Educational Sciences: Theory and Practice*, *12*(2), 1190–1194.

Pao, L. S. (2014). *Effects of keyword generation and peer collaboration on metacomprehension accuracy in middle school students* [Columbia University]. https://doi.org/10.7916/D8HX19TV

Parkinson, M. M., Dinsmore, D. L., & Alexander, P. A. (2010). Calibrating calibration: Towards conceptual clarity and agreement in calculation. *Annual Meeting of the American Educational Research Association*.

Pintrich, P. R., Smith, D., Garcia, T., & Mckeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire* [Technical Report # 91-B-004.]. School of Education, University of Michigan.

Pressley, M., & Gaskins, I. W. (2006). Metacognitively Competent Reading Comprehension Is Constructively Responsive Reading: How Can Such Reading Be Developed in Students? *Metacognition and Learning*, *1*, 99–113. https://doi.org/10.1007/s11409-006-7263-7

Pressley, M., & Ghatala, E. S. (1990). Self-regulated learning: Monitoring learning from text. *Educ. Psychol*, *25*, 19–33.

Prins, F. J., Veenman, M. V. J., & Elshout, J. J. (2006). The impact of intellectual ability and metacognition on learning: New support for the threshold of problematicity theory. *Learning and Instruction*, *16*(4), 374–387.

Reynolds, R. E. (1992). Selective attention and prose learning: Theoretical and empirical research. *Educ. Psychol. Rev*, *4*, 345–391.

Roschelle, J., & Teasley, S. D. (1995). The Construction of Shared Knowledge in Collaborative Problem Solving. In *Computer Supported Collaborative Learning* (pp. 69–97). https://doi.org/10.1007/978-3-642-85098-1_5

Rudolph, J., Greiff, S., Strobel, A., & Preckel, F. (2018a). Understanding the Link Between Need for Cognition and Complex Problem Solving. *Contemporary Educational Psy-Chology*. https://doi.org/10.1016/j.cedpsych.2018.08.001

Rudolph, J., Greiff, S., Strobel, A., & Preckel, F. (2018b). Understanding the Link Between Need for Cognition and Complex Problem Solving. *Contemporary Educational Psy-Chology*. https://doi.org/10.1016/j.cedpsych.2018.08.001

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

Salonen, P., Vauras, M., & Efklides, A. (2005). Social interaction-what can it tell us about metacognition and coregulation in learning? *European Psychologist*, *10*(3), 199–208.

Schneider, W., & Pressley, M. (1989). *Memory Development Between 2 and 20*. Springer-Verlag.

Schoenfeld, A. H., & Herrmann, D. J. (1982). Problem perception and knowledge structure in expert and novice mathematical problem solvers. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *8*, 484–494.

Schraw, & Dennison. (1994a). Assessing metacognitive awareness. *Contem-Porary Educational Psychology*, *19*, 460–475. https://doi.org/10.1006/ceps.1994.1033

Schraw, G. (1995). Measures of feeling-of-knowing accuracy: A new look at an old problem. *Applied Cognitive Psychology*, *9*(4), 321–332.

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning*, *4*, 33–45. https://doi.org/10.1007/s11409-008-9031-3.

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, *36*(1–2), 111–139.

Schraw, G., & Dennison, R. S. (1994b). Assessing metacognitive awareness. *Contem-Porary Educational Psychology*, *19*, 460–475. https://doi.org/10.1006/ceps.1994.1033

Schraw, G., Kuch, F., & Gutiérrez, A. P. (2012). Measure for measure: Calibrating ten commonly used calibration scores. *Learning and Instruction*, *24*, 48–57. https://doi.org/10.1016/j.learninstruc.2012.08.007.

Schraw, G., Potenza, M. T., & Nebelsick-Gullet, L. (1993). Constraints on the calibration of performance. *Contemporary Educational Psychology*, *18*(4), 455–463.

Schraw, G., Wise, S. L., & Roos, L. L. (2000). Metacognition and computer-based testing. In G. Schraw & J. C. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 223–260). Buros Institute of Mental Measurement.

Schraw, & Moshman. (1995). Metacognitive theories. *Educational Psychology Review*, *7*(4), 351–371. https://doi.org/10.1007/BF02212307

Serra, M. J., & Metcalfe, J. (2009). 15 Effective Implementation of Metacognition. In *Handbook of metacognition in education* (p. 278).

Shea, G. P., Guzzo, R., & A. (1987). Group effectiveness: What really matters? *Sloan Man-Age. Rev*, *28*, 25–31.

Shin, N., Jonassen, H. D., & McGee, S. (2003). Predictors of well-structured and illstructured problem solving in an astronomy simulation. *Journal of Research in Science Teaching*, *40*(1), 6–33.

Smith. (1982). Learning strategies of mature college learners. *Journal of Reading*, *26*(1), 5–12.

Smith. (1985). Comprehension and comprehension monitoring by experienced readers. *Journal of Reading*, *28*(4), 292–300.

Snijders, T., & Bosker, R. (1999). Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling. *Http://Lst-Iiep.Iiep-Unesco.Org/Cgi-Bin/Wwwi32.Exe/[In=epi-doc1.in]/?T2000=013777/(100).*

Solomon, G. (1984). Television is "easy" and print is "tough": The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, *76*, 647–658.

Sonnleitner, Brunner, M., Greiff, S., Funke, J., Keller, U., Martin, R., & Latour, T. (2012). The Genetics Lab: Acceptance and psychometric characteristics of a computer-based microworld assessing complex problem solving. *Psychological Test and Assessment Modeling*, *54*, 54–72.

Sperling, R. A., Howard, B. C., Staley, R., & DuBois, N. (2004). Metacognition and Self-Regulated Learning Constructs. *Educational Research and Evaluation*, *10*(2), 117–139. https://doi.org/10.1076/edre.10.2.117.27905

Spring, C. (1985). Comprehension and study strategies reported by university freshmen who are good and poor readers. *Instructional Science*, *14*(2), 157–167.

Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence*. https://doi.org/10.1016/j.intell.2015.09.005

Stadler, Matthias, Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence*. https://doi.org/10.1016/j.intell.2015.09.005

Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-Supported Collaborative Learn-ing. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 479–500). Cambridge University Press. https://doi.org/10.1017/CBO9781139519526.029

Stanovich, K. E. (1990). Concepts in developmental theories of reading skill: Cognitive resources, automaticity, and modularity. *Devel. Rev*, *10*, 72–100.

Striking a balance: Socio-emotional processes during argumentation in collaborative learning interaction. (2018). *Learning, Culture and Social Interaction*. https://doi.org/10.1016/j.lcsi.2017.09.003

Strohschneider, S., & güß, D. (1999). The fate of the moros: A cross-cultural exploration of strategies in complex and dynamic decision making". *International Journal of Psy-Chology,Vol*, *35*(4), 235–252.

Student performance in the postsecondary classroom. (n.d.). *The Journal of Experimental*.

Swanson, H. L. (1990). Influence of metacognitive knowledge and aptitude on problem solving. *Journal of Educational Psychology*, *82*(2), 306–314. https://doi.org/10.1037/0022-0663.82.2.306

Tanner, K. D. (2012). Promoting student metacognition. *CBE—Life Sciences Education*, *11*(2), 113–120.

Toy, S. (2007). *Online ill-structured problem-solving strategies and their influence on problem-solving performance* [Doctor of Philosophy, Iowa State University, Digital Repository]. https://doi.org/10.31274/rtd-180813-17116

Ucan, S., & Webb, M. (2015). Social regulation of learning during collaborative inquiry learning in science: How does it emerge and what are its functions? *International Journal of Science Education*, *37*(15), 2503–2532.

Vauras, M., Iiskala, T., Kajamies, A., Kinnunen, R., & Lehtinen, E. (2003). Shared-regulation and motivation of collaborating peers: A case analysis. *Psychologia: An International Journal of Psychology in the Orient*, *46*, 19–37.

Veenman, M. V., Van Hout-Wolters, B. H., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, *1*(1), 3–14.

Ven, A. H., Delbecq, A. L., & Koenig, R. (1976). Determinants of coordination modes with organizations. *Am. Sociol. Rev*, *41*, 322–338. https://doi.org/10.2307/2094477

Volet, S., Vauras, M., Salo, A.-E., & Khosa, D. (2017). Individual contributions in student-led collaborative learning: Insights from two analytical approaches to explain the quality of group outcome. *Learning and Individual Differences*, *53*, 79–92. https://doi.org/10.1016/j.lindif.2016.11.006

Vollmeyer, R., & Rheinberg, F. (1999). Motivation and metacognition when learning a complex system. *European Journal of Psychology of Education*, *14*(4), 541–554. https://doi.org/10.1007/BF03172978

Vollmeyer, R., & Rheinberg, F. (2000). Does motivation affect performance via persistence? *Learning and Instruction*, *10*(4), 293–309. https://doi.org/10.1016/S0959-4752(99)00031-6

Wageman, R. (1995). Interdependence and group effectiveness. *Adm. Sci. Q*, *40*, 145–180. https://doi.org/10.2307/2393703

Wagener, D. (2001). *Psychologische Diagnostik mit komplexen Szenarios [Psychological diagnostics using complex scenarios*. Pabst.

Walczyk, J. J., & Hall, V. C. (1989). Effects of examples and embedded questions on the accuracy of comprehension self-assessments. *Journal of Educational Psychology*, *81*, 435–437.

Weinert, F., & Kluwe, R. (Eds.). (1984). Mysterious mechanisms. In *Metacognition, motivation*.

Williams, C. (2007). Research methods. *Journal of Business & Economics Research (JBER*, *5*(3).

Wilson, J., & Clarke, D. (2004). Towards the modelling of mathematical metacognition. *Mathematics Education Research Journal*, *16*(2), 25–48.

Winter, G. (2000). A comparative discussion of the notion of validity in qualitative and quantitative research. *The Qualitative Report*, *4*(3), 1–14.

Wüstenberg, S., Greiff, S., & Funke, J. (2012). Complex problem solving—More than reasoning? *Intelligence*. https://doi.org/10.1016/j.intell.2011.11.003

Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, *42*(1), 121–130.

## Appendix 1

### 52 Metacognitive Awareness Inventories in a form of survey with Likert Scale

| Statement | Scale |
|---|---|
| | 0 10 20 30 40 50 60 70 80 90 100 |
| 1. I ask myself periodically if I am meeting my goals. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 2. I consider several alternatives to a problem before I answer. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 3. I try to use strategies that have worked in the past. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 4. I pace myself while learning in order to have enough time. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 5. I understand my intellectual strengths and weaknesses. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 6. I think about what I really need to learn before I begin a task | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 7. I know how well I did once I finish a test. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 8. I set specific goals before I begin a task. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 9. I slow down when I encounter important information. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 10. I know what kind of information is most important to learn. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 11. I ask myself if I have considered all options when solving a problem. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 12. I am good at organizing information. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 13. I consciously focus my attention on important information. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 14. I have a specific purpose for each strategy I use. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 15. I learn best when I know something about the topic. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 16. I know what the teacher expects me to learn. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 17. I am good at remembering information. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 18. I use different learning strategies depending on the situation. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 19. I ask myself if there was an easier way to do things after I finish a task. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 20. I have control over how well I learn. | \|-------\|-------\|------\|------\|------\|------\|------\|------\|------\|------\| |
| 21. I periodically review to help me understand important relationships. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 22. I ask myself questions about the material before I begin. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 23. I think of several ways to solve a problem and choose the best one. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |

<table>
<tr><td>24. I summarize what I've learned after I finish.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>25. I ask others for help when I don't understand something.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>26. I can motivate myself to learn when I need to</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>27. I am aware of what strategies I use when I study.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>28. I find myself analyzing the usefulness of strategies while I study.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>29. I use my intellectual strengths to compensate for my weaknesses.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>30. I focus on the meaning and significance of new information.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>31. I create my own examples to make information more meaningful.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>32. I am a good judge of how well I understand something.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>33. I find myself using helpful learning strategies automatically.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>34. I find myself pausing regularly to check my comprehension.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>35. I know when each strategy I use will be most effective.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>36. I ask myself how well I accomplish my goals once I'm finished.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>37. I draw pictures or diagrams to help me understand while learning.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>38. I ask myself if I have considered all options after I solve a problem.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>39. I try to translate new information into my own words.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>40. I change strategies when I fail to understand.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>41. I use the organizational structure of the text to help me learn.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>42. I read instructions carefully before I begin a task.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>43. I ask myself if what I'm reading is related to what I already know.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>44. I reevaluate my assumptions when I get confused.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>45. I organize my time to best accomplish my goals.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>46. I learn more when I am interested in the topic.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>47. I try to break studying down into smaller steps.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>48. I focus on overall meaning rather than specifics.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
<tr><td>49. I ask myself questions about how well I am doing while I am learning something new.</td><td>|-------|-------|-------|------|------|------|------|------|------|------|</td></tr>
</table>

| | |
|---|---|
| 50. I ask myself if I learned as much as I could have once I finish a task. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 51. I stop and go back over new information that is not clear. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |
| 52. I stop and reread when I get confused. | \|-------\|-------\|-------\|------\|------\|------\|------\|------\|------\|------\| |