Computer-Supported Collaborative Learning (2019)

This is a post-peer-review, pre-copyedit version of an article published in the International Journal of Computer-Supported Collaborative Learning. The final authenticated version is available online at: https://doi.org/10.1007/s11412-018-9293-y

Providing different types of group awareness information to guide collaborative learning

Running head: Group Awareness Information

Lenka Schnaubert * Daniel Bodemer

lenka.schnaubert@uni-due.de, bodemer@uni-due.de
University of Duisburg-Essen
Lotharstrasse 65, 47057 Duisburg, Germany

Corresponding Author: Lenka Schnaubert
eMail: lenka.schnaubert@uni-due.de

Abstract Cognitive group awareness tools are a means to guide collaborative learning activities by providing knowledge-related information to the learners. While positive effects of such tools are firmly established, there is no consistency with regard to the awareness information used and a wide range of target concepts exist. However, attempts to compare and integrate the effects of different types of group awareness information are rare. To reduce this gap, our study aims to compare metacognitive and cognitive group awareness information, combining CSCL research and research on metacognition. In our experimental study, 260 university students discussed assumptions on blood-sugar regulation and diabetes mellitus in dyads. We tested the effects of providing cognitive group awareness information on the learners' assumptions (factor 1) and metacognitive group awareness information on their confidence (factor 2) on individual metacognitive and cognitive outcome measures and on the learners' regulation of the collaborative process, i.e., the selection of discussion topics based on confidence in knowledge (confidence-based regulation) and based on agreement regarding assumptions (conflict-based regulation). We found that visualizing information strongly impacts joint regulation and that learners seem to integrate the information provided to steer their learning. However, while the learners gained knowledge and confidence during collaboration, providing group awareness information did not have the expected impact on learning outcomes. Reasons and implications of these results in light of previous research on metacognition and group awareness are discussed.

Keywords Computer-supported collaborative Learning * Group Awareness * Guidance * Metacognition * Self-regulated Learning

Introduction

Collaborative learning has great potential to strengthen learners' content-related and meta-skills. However, learners face many challenges when attempting to learn together, especially concerning communication and coordinating their activities (e.g., G. Erkens, Jaspers, Prangsma, & Kanselaar, 2005; Janssen, Erkens, & Kanselaar, 2007). Thus, guidance is considered to be an important part of (computer-supported) collaborative learning (e.g., Fischer, Kollar, Stegmann, & Wecker, 2013). Research on implicit guidance is focusing on approaches that support learners' self-regulation attempts and foster learners' agency (cf. Hacker, Dunlosky, & Graesser, 2009). Such approaches provide relevant information for learners without giving them explicit structure or instructions, leaving the locus of control with the learners and building directly on individual skills (Hesse, 2007; Janssen & Bodemer, 2013; Miller & Hadwin, 2015).

One prominent way to implicitly guide collaborative learning processes is the provision of knowledge-related group awareness information (cf. Janssen & Bodemer, 2013). Group awareness (GA) is the state of being informed about relevant aspects of group members or the group as a whole (Bodemer & Dehler, 2011; Bodemer, Janssen, & Schnaubert, 2018), for example their knowledge and skills. Collaborating learners need an awareness of such aspects to effectively steer the collaboration process and adjust it to the needs of the group (Bodemer & Dehler, 2011; Fransen, Kirschner, & Erkens, 2011; Soller, Martínez, Jermann, & Muehlenbrock, 2005). If information about the cognitions of other learners is missing, learners may tend to overestimate similarities (Nickerson, 1999) and might thus fail to detect relevant differences in knowledge and/or opinions. Although such information can be provided within the learning situation by using specifically designed tools (GA tools) that support learners' formation of GA (cf. Bodemer et al., 2018), there is a lack of research investigating how different types of knowledge-related GA information within such tools guide learning processes and how this affects learning outcomes.

Group awareness tools

GA tools facilitate GA by providing learners with relevant information about their learning partners. While GA tools used for computer-supported collaborative learning (CSCL) may target various types of learner-related information (including cognitive and/or social variables; cf. Janssen & Bodemer, 2013), they usually process the information in a three step manner: they assess relevant information, transform it and feed it back to the learners, usually by visualizing it in an adequate way (cf. Buder & Bodemer, 2008). All three steps may highly depend on technological support to assess the information (e.g., using computer-based questionnaires or logfile data), transform the information (e.g., by using specific algorithms to condense and thus pre-interpret the information), and visualize it (e.g., by converting the information into a graphical representation within the learning environment).

What sets these tools apart from other tools based on learning analytics is the target audience. While other educational tools relying on learner data often feed information to educators or adaptive systems (e.g., via teacher dashboards, adaptive learning environments, computer-based pedagogical agents), GA tools feed the information back to the learners themselves, meaning that data subjects are identical to data clients (cf. Greller & Drachsler, 2012). This is a vital distinction, because it requires the information to be adapted to characteristics of the learners as target audience as opposed to educators or tool designers. Limitations of the learners as data clients thus put restrictions on the usage of learning analytics as it is assumed that learners are often not competent enough to learn from learning analytics reports unsupported (Drachsler & Greller, 2012). As a consequence, the provision of GA information has to be tailed to the needs of the learners. Several researchers thus point out, that awareness information has to be perceived as useful (e.g., Janssen, Erkens, & Kirschner, 2011; Nova, Wehrle, Goslin, Bourquin, & Dillenbourg, 2007) and authentic (e.g., Engelmann, Dehler, Bodemer, & Buder, 2009), easy to understand and interpret (e.g., Bodemer, 2011; Dehler, Bodemer, Buder, & Hesse, 2011), and saliently displayed (e.g., Bodemer & Dehler, 2011). This has an impact on all three data processing steps, as it relates to the information itself (i.e., perceived usefulness), the assessment and transformation of the information (i.e., authenticity and interpretability) as well as the presentation (i.e., salience).

GA tools focusing on knowledge-related information (often called "cognitive GA tools", e.g., Bodemer et al., 2018) foster collaborative learning processes by making learners aware of each other's individual or their common cognitive status or processes (Bodemer & Dehler, 2011). These tools may benefit collaborative learning in several ways. By systematically processing knowledge-related information externally (assessing, transforming, presenting), they facilitate the natural formation of GA by adding an external reference (Engelmann et al., 2009) and thus relieve the learners of effortfully extracting relevant information themselves. This can be of vital importance, especially if germane learning processes take up most of the available cognitive resources. Another key function of these tools is to structure learning discourses (cf. Bodemer et al., 2018). By focusing on specific, pre-defined GA information and processing it in a specific way, GA tools offer an interpretation of the collaborative situation and thereby suggest specific courses of action beneficial to learning (via informational and representational guidance; cf. Bodemer, 2011). For example, the information provided may ease the identification of individual (or group) needs or conflicting viewpoints (Engelmann et al., 2009). If the GA information is linked to the learning material, it can focus collaborating learners on aspects of the learning material that need further attention (Bodemer et al., 2018; Bodemer & Scholvien, 2014), helping them to structure and coordinate mutual learning processes (Clark & Brennan, 1991). Further, the availability of such information may trigger beneficial collaboration processes like exchanging and explaining relevant knowledge (Dehler, Bodemer, Buder, & Hesse, 2009; Dehler et al., 2011), argumentation or elaboration (Buder & Bodemer, 2008; Dehler, Bodemer, & Buder, 2007). These processes are beneficial for learning and although they are more likely to occur during collaboration than in individual learning settings, they do not necessarily occur spontaneously (Dillenbourg, Järvelä, & Fischer, 2009; King, 2007).

GA tools may support very different processes relevant for learning. Consequently, there is a multitude of tools that provide very different kinds of information assessed in very different ways (for an overview of different tools see Janssen & Bodemer, 2013). While overall, beneficial effects have been firmly established, one of the challenges of the research field is to systematically explore how these tools foster learning (Buder, 2011) and what tool features are responsible for the effects to ultimately provide efficient and precise support for learners (Bodemer et al., 2018). Thus, our research aims at looking into a distinct feature of GA tools used for CSCL namely the type of learner-related information they portray to learn more about the guidance mechanisms involved to inform researchers, tool designers, and educators alike.

Cognitive and metacognitive information in collaboration

As stated above, GA tools that aim for guidance effects vary greatly in what information they present and thus the learning material they may bring to the learners' attention (and the processes they may trigger). *Cognitive group awareness tools* provide knowledge-related information of the group, e.g., how much learners know (M. Erkens, Bodemer, & Hoppe, 2016; Sangin, Molinari, Nüssli, & Dillenbourg, 2008, 2011), what they think (Bodemer, 2011; Engelmann & Hesse, 2010, 2011; Gijlers, 2005; Gijlers, Saab, van Joolingen, de Jong, & van Hout-Wolters, 2009), or how they judge their knowledge (Dehler et al., 2009, 2011). Research in this area usually adopts an inclusive conception of the term "cognitive" and rarely explicitly differentiates between different types of cognitive information like information about the content of knowledge and information about the learners' metacognitive evaluation of said knowledge. In this line of research, it is often assumed that asking learners to judge their own knowledge is just a short-cut for assessing their actual knowledge. However, in metacognition research, such self-evaluations are seen to have additional benefits and may represent the learner's perspective and thus serve as the basis for their self-regulatory actions. The term "metacognition" originated in the 1970s and has been conceptually described as "thinking about thinking" (e.g., Flavell, 1979). In more recent work, the term comprises of various concepts and processes relating to monitoring, controlling and/or regulating learning processes (e.g., Dinsmore, Alexander, & Loughlin, 2008) including, for example, monitoring knowledge and knowledge acquisition, and planning, controlling and evaluating learning processes.

Within the metamemory framework terminology of Nelson and Narens (1990), information on the content of learners' knowledge is information on the object level (cognitive), while information on learners' self-evaluation of said knowledge constitutes information on the meta-level (metacognitive). Since the term "cognitive" falls short of such a differentiation, we will use the broader term "knowledge-related" for these kinds of information and refer to "cognitive information" when talking about object-level information and the term "metacognitive information" when addressing meta-level information. Using a metacognition framework on GA tools may benefit CSCL research as it allows to differentiate tool effects within collaborative learning by drawing on metacognition theory and research.

Effects of cognitive information: identifying conflict

As stated above, some knowledge-related GA tools provide information on the content of cognitions of group members, such as learners' opinions or assumptions (Bodemer, 2011; Gijlers et al., 2009). If these are presented next to each other and in a similar format, they may foster comparison processes and thus promote the detection of conflicting assumptions on the topic to be learned between learners within a collaborating group (Scholvien & Bodemer, 2013). If learners are not aware of such epistemic conflicts, these conflicts may hamper progress and a shared mental representation of the material to be learned. Additionally, conflicting assumptions are an essential drive to cognitive development (e.g., Levine, Resnick, & Higgins, 1993; Mugny & Doise, 1978; see also Doise & Mugny, 1984). They provide the need to build a consensus and thus, they do not only provide the opportunity to adjust faulty assumptions, but they also provide an opportunity to elaborate on the content while discussing the positions from different perspectives and resolving the conflict in a beneficial way (Dillenbourg & Hong, 2008; Gijlers et al., 2009; D. W. Johnson & Johnson, 2009b). Conflicting assumptions are widely assumed to produce uncertainty about the correctness of assumptions (e.g., Buchs, Butera, Mugny, & Darnon, 2004; Crano & Prislin, 2006; Fraundorf & Benjamin, 2016; Koriat, Adiv, & Schwarz, 2015), which may also trigger beneficial learning processes. However, uncertainty may prevail if learners do not get the chance to resolve these conflicts sufficiently (Schnaubert & Bodemer, 2016).

Empirically, conflicting assumptions have been found to attract attention in collaborative settings and are discussed more frequently than matching (i.e., congruent) assumptions (Bodemer, 2011). Further, they may trigger beneficial evaluation processes (Buchs et al., 2004; Doise, Mugny, & Perret-Clermont, 1975; D. W. Johnson, Johnson, & Tjosvold, 2000; R. Johnson, Brooker, Stutzman, Hultman, & Johnson, 1985; Mugny, Butera, Sanchez-Mazas, & Perez, 1995; Mugny & Doise, 1978) as well as the search for information (Buchs et al., 2004; Lowry & Johnson, 1981; Schnaubert & Bodemer, 2016) or coordination efforts of the learning process (Mugny & Doise, 1978). In sum, epistemic conflicts can be a driving force during collaborative learning. To benefit from these mechanisms, learners must become aware of the conflicts, and technology (i.e. GA tools) can help to process the necessary information in a beneficial way. Empirically, cognitive GA tools have been found to draw attention towards conflicts (e.g., Bodemer, 2011; Schnaubert & Bodemer, 2016) and have supported conflict resolution (Bodemer, 2011).

Effects of metacognitive information: identifying perceived lacks of knowledge

Other tools do not provide information on the content of learners' cognitions, but on the presence or absence (or degrees) of knowledge. This provides a basis for clearing up lacks of knowledge and has been found to be an effective means to support learning (e.g., Dehler et al., 2009, 2011; Sangin et al., 2011). Assessing knowledge from the learners' perspective allows for authentically capturing the learners' perceived need for information (Engelmann et al., 2009). Such metacognitive information has been used as a target concept for GA tools. For example, Dehler and colleagues (Dehler et al., 2009, 2011) conducted a series of studies providing learners with information on the perceived understanding of content via text comprehension ratings. Such information informs the group about (perceived) lacks of knowledge and knowledge distributions and have been found to guide communication (Dehler et al., 2009, 2011). In this research, comprehension ratings were interpreted as providing information on the presence or absence of knowledge regarding a specific topic, however, from a metacognitive point of view, these ratings provide information about the learners' metacognitions, i.e., metacomprehension.

Within metacognition research, metacognitive monitoring is seen as an essential drive for self-regulation activities (e.g., Nelson & Narens, 1990; Thiede & Dunlosky, 1999; Winne & Hadwin, 1998) as it is thought to provide internal feedback that learners can use to steer their learning processes (Butler & Winne, 1995). Empirical research in this area has repeatedly shown that the result of metacognitive monitoring of learning processes and outcomes is (causally) related to how learners control their learning (e.g., Efklides, Samara, & Petropoulou, 1999; Kornell & Metcalfe, 2006; Metcalfe & Finn, 2008; Nelson, Dunlosky, Graf, & Narens, 1994; Nelson & Leonesio, 1988; Son & Metcalfe, 2000), i.e., what they choose to study and for how long. Such metacognitive regulation can have positive effects on learning outcomes (e.g., Kornell & Metcalfe, 2006; Nelson et al., 1994), especially if monitoring is accurate and thus indicative for performance (Dunlosky & Rawson, 2012; Thiede, 1999; Thiede, Anderson, & Therriault, 2003). Further, monitoring-based (metacognitive) regulation may be fostered by visualizing monitoring outcomes (e.g., response confidence ratings, Schnaubert & Bodemer, 2017). In turn, how learners control their learning also affects monitoring (Koriat, Ma'ayan, & Nussinson, 2006). Thus, monitoring, regulation, and performance are inherently linked (cf. Special Issue Koriat, 2012).

One core metacognitive concept is response confidence (Dunlosky & Metcalfe, 2009; Nelson & Narens, 1990). Drawn from metamemory research, response confidence is an evaluation of preceding performance, i.e., performance monitoring (e.g., Dunlosky & Hertzog, 2000; Hines, Touron, & Hertzog, 2009). Thus, it takes specific test experience into account and is connected to specific assumptions about a topic rather than an overall state of learning. Empirically, it is connected to re-study decisions (e.g., Hines et al., 2009; Schnaubert & Bodemer, 2017), and feedback processing (e.g., Fazio & Marsh, 2009; Kulhavy & Stock, 1989). Moreover, it is diagnostic of performance in some circumstances (Koriat et al., 2006; Maki, 1998b, 1998a). Theoretically, confidence in own assumptions may be viewed as an essential part of knowledge itself (Hunt, 2003) and even objectively correct assumptions require a minimum of confidence to be usable in practice (i.e., to guide decision-making or behavior; cf. Leclercq & Poumay, 2004). Accordingly, confidence in test responses has also been used as additional information in knowledge assessment (confidence marking; Leclercq, 1983, 1993). Since confidence in knowledge may be directly linked to the content of knowledge (i.e., specific assumptions), confidence in responses seem to be particularly suitable to represent learners' metacognitions about their knowledge in a collaborative learning context.

Overall, within metacognition research, metacognitive monitoring and monitoring-based regulation play a crucial role in self-regulating individual learning. However, metacognitive regulation may also be affected by information on other learners' knowledge (Schnaubert & Bodemer, 2016). Within a social scenario, externalizing metacognitive evaluations of knowledge allows learners to intentionally disclose gaps in knowledge (from their perspective) and may thus be used strategically to communicate a need for information. Conversely, learners may detect gaps in their learning partners' knowledge (e.g., by viewing GA information) and use the information to support the partners' learning processes by offering and/or adapting help (Dehler et al., 2007, 2009, 2011). Adapting utterances to the properties of the listener (audience design) is vital for effective communication (Clark & Murphy, 1982) and inherently linked to the common ground of learners interacting (Clark & Brennan, 1991). While there is an increasing interest in regulatory processes within collaborating groups (e.g., socially shared regulation; e.g., Hurme, Palonen, & Järvelä, 2006; Iiskala, Vauras, Lehtinen, & Salonen, 2011; Järvelä & Hadwin, 2013; Järvelä et al., 2016; for an overview see Panadero & Järvelä, 2015), the role of the availability of metacognitive information within CSCL has received only scarce attention so far, although there are notable exceptions (e.g., Järvelä et al., 2015; however, these refer to regulatory activities like planning and task perception rather than metacognitive evaluations of knowledge and memory). Thus, despite research showing that visualizing information on (actual or perceived) lacks of knowledge draws learners' attention and can guide learning and communication processes (e.g., asking questions or providing spontaneous explanations, e.g., Bodemer, 2011; Dehler et al., 2011), there is a lack of research utilizing metacognitive GA information within the context of CSCL and explicitly linking it to metacognition research.

Interaction of cognitive and metacognitive information

While assessing the effects of both cognitive and metacognitive GA information on collaborative learning is firmly based on prior empirical evidence, the two types of information are not independent. There are various ways in which the information of multiple learners may interact and influence each other. In the following, we will briefly describe how other learners' cognitions may affect our own cognitions and metacognitive evaluations, how our own metacognitions may affect how we process information about other learners, and how others' metacognitive evaluations may affect how much credit we give them as a source of information.

First, information on others' cognitions may be used as feedback on one's own cognitions and thus change not only the cognitions themselves, but also our metacognitive evaluation of them. For example, if a learner's assumption is supported by congruent assumptions of other learners, he/she may be ensured of his/her position. Conversely, firm beliefs may be rattled if learning partners disagree with the individual's assumption, leading to uncertainties (for an overview on social influence affecting individual confidence see Koriat et al., 2015).

However, our metacognitions (such as confidence in our knowledge and assumptions) may also affect how we process incoming information. Within research on cognitive feedback, confidence takes an important role and research indicates that feedback on errors committed with high-confidence is treated differently than feedback on errors committed with low-confidence (Hancock, Stock, & Kulhavy, 1992; Kulhavy & Stock, 1989). For example, feedback messages on high-confidence errors are studied longer (Fazio & Marsh, 2009; Kulhavy & Stock, 1989) and high-confidence errors are corrected more often than low-confidence errors (Butterfield & Metcalfe, 2001, 2006; Metcalfe & Finn, 2011). Similar effects might apply to learning partners disagreeing with low- and high-confidence assumptions, even though a learning partner is not an indisputable source like expert feedback usually is. Thus, how learners evaluate their learning partners' cognitions and competences becomes increasingly relevant and one indicator to use are the partners' own metacognitions.

Consequently, another way cognitive and metacognitive information of collaborating learners interact is that others' metacognitive evaluations may affect how we evaluate their cognitions as well. Metacognitive evaluations of distinct cognitions (like chunks of knowledge or assumptions) may be used to qualify these cognitions (Hunt, 2003; Leclercq & Poumay, 2004) since they may be treated as an indicator for comprehension (Kulhavy, Stock, Hancock, Swindell, & Hammrich, 1990), which may help learners to better understand their peer's position. For example, confidence in assumptions can have a great impact on how these assumptions are perceived. Research on source reliability and credibility has found that confidence is a major factor in estimating whether someone is a reliable source of information (e.g., Tenney, Small, Kondrad, Jaswal, & Spellman, 2011) and thus confidence is used to evaluate knowledge and competence of others as well as accuracy of the information provided (Price & Stone, 2004; Yates, Price, Lee, & Ramirez, 1996).

In sum, confidence in one's own assumptions yields information relevant to individual learning processes that can be used by other learners to judge the learning partners' cognitive status and support the joint regulation of the collaborative learning process. Since the metacognitive and cognitive level are inherently intertwined, the provision of information on both might thus have a distinct effect on learning. While it would seem logical that these two information types might interact in steering learning processes (cf. Schnaubert & Bodemer, 2016), it may also be the case that learners focus on just one aspect because they regard one type of information as more important or because they need to reduce the strain posed on the cognitive system. Thus, while mainly individual-focused research suggests interaction effects between information on the cognitive and metacognitive level, group situations are *per se* much more complex (cf. Dillenbourg & Bétrancourt, 2006) and simplifying courses of action may also be a strategy to handle provided information.

The following study connects individual-focused research on metacognition with research on CSCL, by experimentally pursuing whether metacognitive and cognitive GA information in GA tools interact in guiding collaborative learning processes and fostering learning outcomes or if they are regarded independently. Analyzing the effects these types of information have on learners when integrated in GA tools will help to improve GA tools and adapt them towards the specific goals of educators and/or tool designers.

Research questions and hypotheses

Ultimately, this study aims at combining two very different types of knowledge-related GA information within GA tools: one providing information on the learners' assumptions capturing learners' cognitions about the topic (cognitive information) and one providing information on the learners' confidence in their assumptions capturing metacognitive evaluations of said knowledge (metacognitive information). Individually they may trigger very different mechanisms: information about specific assumptions may be compared between learners and can evoke socio-cognitive conflicts if assumptions differ and can thus reveal a need for clarification. If assumptions match, learners may see their assumptions validated and abstain from further engagement. Information on learners' metacognitive states may help identify (perceived) lacks of knowledge within the group and thus a need for further engagement with a specific topic. Additionally, recognizing when learners are confident about their knowledge may also help to identify available resources. In combination, both types of information may be used to easily link cognitive and metacognitive information – this may lead to different foci on conflicting assumptions or lacks of knowledge and thus alternative behavioral approaches.

In accordance with research on guidance mechanisms of GA tools, we assume that metacognitive GA information will lead to a focus on needs for information and thus to the primary discussion of aspects learners in the group are unsure about – with low confidence pointing towards a need for clarification within the group. In metacognition research, this type of selection based on metacognitive evaluation of the state of learning is occasionally referred to as "metacognitive regulation" (Thiede, 1999; Thiede et al., 2003). Accordingly, collaborative metacognitive regulation on a group level may be addressed by investigating if metacognitive evaluations on the subjects' knowledge of specific topics guide discussion of said topics. Because research on GA suggests that providing metacognitive information guides learning processes and empirical studies have found that in individual (Schnaubert & Bodemer, 2017) and pseudo-collaborative settings (Schnaubert & Bodemer, 2016) the provision of information on metacognitive confidence ratings may foster metacognitive (i.e., confidence-based) regulation, we hypothesize that providing metacognitive GA information on confidence ratings will foster a structured approach and collaborative metacognitive regulation within groups of learners (H1a) as well. It is further hypothesized that this focus on uncertainties and insufficiently learned topics leads to more accurate knowledge (H2a). Although prior studies using confidence ratings in individual settings have not found this effect on learning outcomes especially if there is low overlap between metacognitive self-evaluations and actual performance (Schnaubert & Bodemer, 2016, 2017), we argue that a learning partner may be able to support individual knowledge construction in the case of uncertainties. Further, uncertainties always indicate a lack of knowledge that needs to be addressed to gain usable knowledge. Explicitly addressing uncertainties should thus give learners the opportunity to clear up uncertainties. Accordingly, we hypothesize higher confidence levels if learners address uncertainties (H3a).

As opposed to metacognitive information, cognitive information has no inherent standard and may thus only be evaluated by measuring it against an external standard. This standard may be a differing opinion of a learning partner. Comparing one's own to a partner's assumption may thus lead to cognitive conflicts and in consequence also a need for clarification on a group level. By providing easily comparable information on learners' assumptions in a GA tool, we assume to foster the identification of such conflicts. Since research on individual learners (Schnaubert & Bodemer, 2016) as well as on dyads of learners (Bodemer, 2011) has shown that learners focus on conflicting assumptions if such information is provided, we assume this will be the case here as well. We hypothesize that learners regulate their learning more strongly based on conflicts (collaborative conflict-based regulation; comparable to metacognitive regulation with the presence or absence of conflicting assumptions as driving force) if cognitive information is provided than if this information is not provided (H1b). Since socio-cognitive conflicts have repeatedly shown to be beneficial for learning (cf. D. W. Johnson & Johnson, 2009a), we further hypothesize that this focus on conflicting assumptions within a collaborative setting fosters learning (H2b). Discussing and resolving conflicts may thus also lead to firmly established knowledge and thus to high confidence levels regarding this knowledge. Focusing on a lack of consensus apparent in conflicts, however, may also unsettle learners and make them doubt their knowledge (cf. Crano & Prislin, 2006; Koriat et al., 2015). If conflicts cannot be completely resolved, discussing them may thus also foster uncertainties. Consequently, we want to know if and how cognitive information affects confidence levels of the learners interacting, but we abstain from formulating specific hypotheses (RQ1).

Providing cognitive and metacognitive GA information within a GA tool may have different effects on CSCL. Theoretically, learners may use both types of information separately to structure their learning or focus on one type of information at a time, ignoring the other. On the other hand, learners might integrate the two kinds of information by weighing cognitive information with metacognitive information to prioritize and steer their learning process more sophisticatedly. Due to the lack of evidence regarding collaborative learning supported by both cognitive and metacognitive GA information, we abstain from being explicit in our assumptions here, but rather aim at exploring if and how groups of learners deal with the combination of both kinds of information.

Methods

Design and sample

To answer our research questions, we conducted a study with $N = 130$ dyads of students (260 students in total). They were mainly first semester university students (62.7%) enrolled in a Bachelors' course on Applied Cognitive and Media Science at a German University. 197 students were female (76%), 63 (24%) male. The mean age was 21.00 years ($SD = 2.69$). 89.2% of the participants judged their knowledge on blood sugar regulation to be low or rather low (values of 0 to 2 on a scale from 0 to 5; $M = 1.11$, $SD = 1.04$), 90.8% judged their knowledge on diabetes mellitus to be low or rather low (analogous scale, $M = 0.97$, $SD = 1.02$). Their interest in the topic was somewhat higher, with only 41.9% claiming a low or rather low interest in blood sugar regulation (values of 0 to 2 on a scale from 0 to 5; $M = 2.72$, $SD = 1.14$) and 40.8% in diabetes mellitus (analogous scale, $M = 2.72$, $SD = 1.16$).

The students could sign up together as a dyad or independently. Each dyad was randomly assigned to one of four research conditions in a 2x2 between-dyad design. Additionally, some dependent variables were assessed repeatedly, so the design includes the within-dyad factor "time of assessment" (pre and post intervention).

The learners conducted the first part of the experiment individually. Then they came together on a multi-touch table top computer for the collaborative phase of the experiment. In this phase, we manipulated the independent variables by varying (1) the display of cognitive GA information on assumptions (displayed: cGAI+ / not displayed: cGAI-) and (2) the display of metacognitive GA information on confidence ratings (displayed: mGAI+ / not displayed: mGAI-). This resulted in four between-dyad experimental conditions: dyads with no GA information provided (mGAI-/cGAI-), dyads with only metacognitive GA information provided (mGAI+/cGAI-), dyads with only cognitive GA information provided (mGAI-/cGAI+) and dyads with both metacognitive and cognitive GA information provided (mGAI+/cGAI+). After collaboration, there was another individual phase where we further assessed dependent variables.

The dependent variables varied not only with respect to number of measurements, but also with respect to level. While individual variables like knowledge or confidence in answers were assessed at an individual level, collaborative process data was assessed on dyadic level. We conducted intra-class-correlations (ICC; cf. Shrout & Fleiss, 1979) based on a single-rating, absolute-agreement, one-way random-effects model to check if data assessed individually was interdependent within dyads and needed to be analyzed on a dyadic level.

Procedure

Learners were invited to take part in the experiments in dyads. After welcoming, introduction and declaration of consent, they conducted the first part of the experiment separately on desktop computers. For the second part, they were asked to collaborate on a multi-touch table top computer. The third part was again conducted separately on the desktop computers (cf. Technical setup section).

An instructor was with the participants throughout the experiment, but did not interfere except to welcome the participants at the beginning and give them a general introduction, to introduce them to the collaboration phase, and to reward them and see them off at the end of the study. Otherwise they interfered only if problems occurred. All instructors acted according to a pre-defined script and were trained by the principle researcher in advance.

First, the instructor started a computer program, which gave all further instructions. The participants first read an introduction, then filled out a demographics questionnaire including questions about their prior knowledge about and interest in the topics blood sugar regulation and diabetes mellitus. Then they each read an assigned learning text on this topic for up to 15 minutes after which they were automatically redirected to the next page. For the last 5 minutes, a countdown was visible to allow the participants to adjust to the available time. They were able to terminate the reading process after a minimum of 10 minutes. Afterwards, they were introduced to learning tasks in the form of true-false questions with three example training items about geography, music and paleontology. They then answered 16 learning tasks including confidence ratings on the topic of blood sugar regulation and diabetes mellitus (t1). When both learners had finished with this task, they were asked by the instructor to come to the multi-touch table top computer. Here, the instructor gave a brief introduction explaining the functions, functionalities and visualizations and handling of the program by using the three training tasks and random text after a fixed script. The instructor gave the participants time to try out the features and answered questions regarding program usage. After they got acquainted with the program, the instructor started the actual collaboration phase, which lasted 16 minutes after which the program shut down automatically. The participants were then led back to their individual computers and resumed the experiment individually. They filled out a questionnaire about their structuring of the collaboration phase, answered the learning tasks again from scratch including confidence ratings (t2) and finished off with a knowledge test. They were thanked and each either rewarded 12 Euros or course credits. The whole experiment lasted 75 to 90 minutes. The procedure can be viewed in Figure 1. The collaboration was video recorded with the camera pointed at the table top, capturing gestures and voice, but not the faces of the participants. Video data was used to clear up (rare) issues with the log data assessed on the table top computer (cf., Logging of collaboration data section).

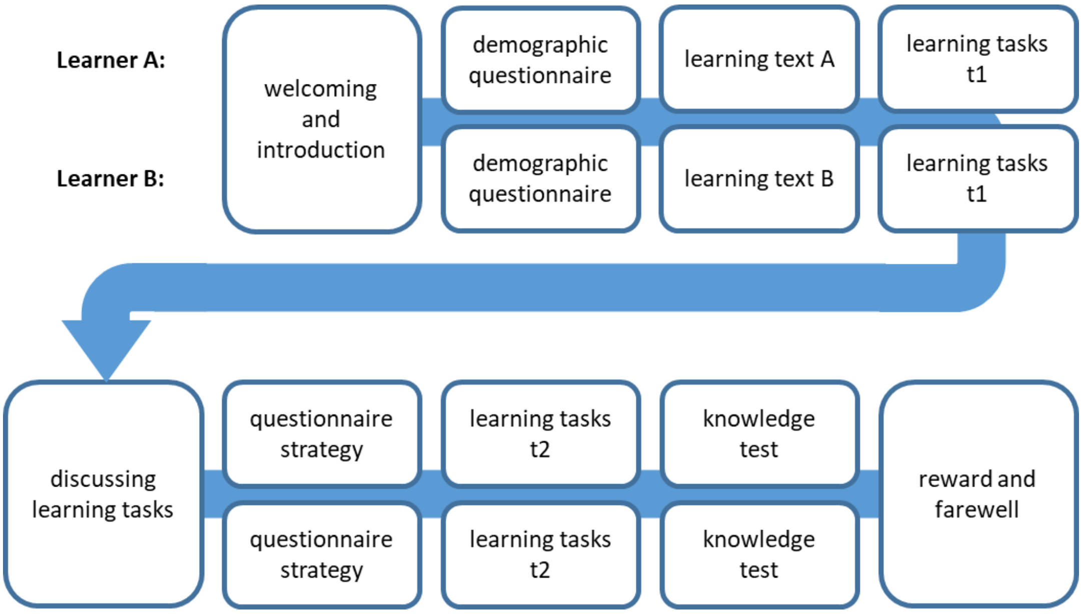

Figure 1: Experimental procedure

Technical setup

During the individual phases, learners worked on individual computers; learners were separated by blinds. Talking between participants was not permitted during this stage. All texts and questionnaires were presented in a pre-defined order using HTML pages and CSS files. The MediaLab v2012 software (Jarvis, 2012) was used to run the experiment and sequence the pages. Client data was saved into an open-source database management system (Apache CouchDB™ v1.5; Apache Software Foundation, 2013) on a central server. The server data was accessible by the multi-touch table top to allow data from the individual phase to be transferred for the collaboration phase.

During the collaborative phase, learning partners worked together on a Samsung SUR40 multi-touch table top computer. The learning environment was programmed using an object-oriented programming language supporting touch events (C#). The client was hooked to the database to access the data from the individual phase and to save data during the collaboration. Before the start of the collaboration, GA information assessed in the individual phase was transferred to the table top. According to experimental condition, data was transformed to allow for the different data visualizations during collaboration (cf. Treatment section).

Material

The material consisted of two texts (one for each learning partner within a dyad), 16 learning tasks and 32 knowledge test items regarding blood sugar regulation and diabetes mellitus, which built on material used in a study on the same topic by Schnaubert and Bodemer (2017). Additionally, we developed a questionnaire asking the learners how they structured their collaborative learning process.

The texts were both 1208 words long and contained ten key paragraphs plus two short introductory paragraphs. The texts had four identical key paragraphs and differed with regard to the six others. For example, one text (A) had a focus on diabetes type 1 and the other (B) on diabetes type 2. Also, text A contained paragraphs explaining the risk and treatment of hyperglycemia, while text B focused on hypoglycemia.

Each of the 16 learning tasks referred to information from exactly one paragraph in any one of the texts and each paragraph was represented by exactly one learning task. Each task consisted of a statement that had to be judged by the participants individually to be true or false (true/false statement; cf. Figure 2). While answering the question by clicking into a respective filed, a pop-up asked learners to state on a binary scale how confident they were, that their answer was correct (confident vs. not confident; cf. Figure 2). A binary scale was chosen to keep the information (and representation) as simple as possible. As argued before, for learners to benefit from GA information, it has to be easily accessible and understandable to guide the learning process without costs with regard to mental load (especially when including multiple types of GA information). Thus, the information needs to be presented in a way that learners understand instantly and fosters comparison processes between learners to detect relevant patterns (cf. Bodemer, 2011), but also between items to support between-items selection processes (cf. Ariel, Dunlosky, & Bailey, 2009).

Texts and tasks were first worked through individually, before learners came together to work on the tasks collaboratively.

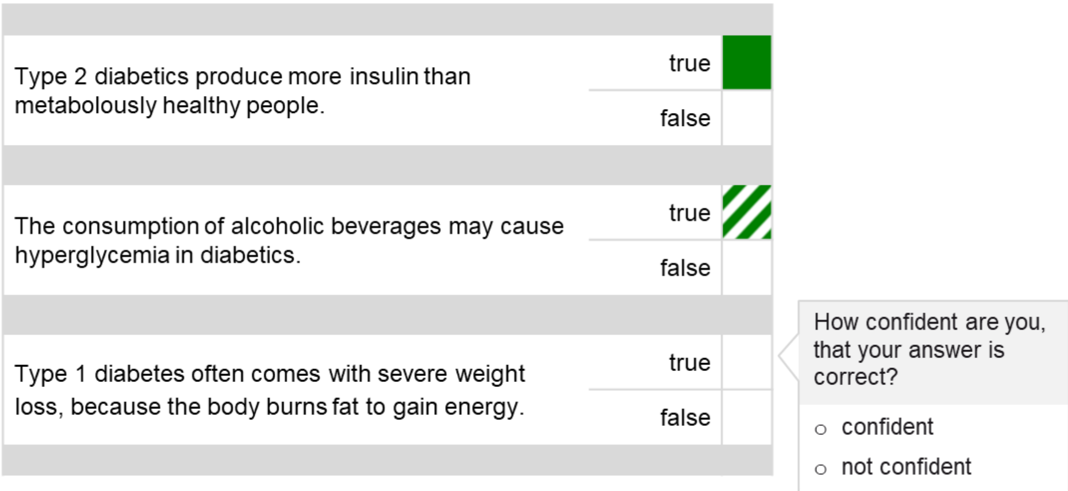

Figure 2: Examples of translated items with answers and confident ratings (confident = full green; not confident = hatched white-green), individual phase

The knowledge test consisted of 32 items – two for each paragraph in the texts and thus learning task. One assessed the knowledge that had also been assessed in the learning tasks more elaborately and one asked for more elaborate information to assess transfer. Again, the tasks had been adapted from the ones used by Schnaubert and Bodemer (2017). Each task was accompanied by a response confidence question (How confident are you, that you solved this task correctly?) on a six-point equidistant ordered response scale ranging from "not confident at all" [0] to "absolutely confident" [5].

We also developed a questionnaire that asked the learners how they structured their learning process. It consisted of two parts. Part one asked the learners to specify how true five statements were with regard to their collaborative learning phase on a six-point equidistant ordered response scales ranging from "not true at all" [0] to "completely true" [5]. Two open-ended questions were included to elaborate on or specify the answers to statements three and five (cf. Table 1). Part two consisted of two questions asking specifically whether they based their selection of tasks on their (un-)certainties (question 1) and whether they based it on their (dis-)agreement (question 2) with an additional question and an open field to include further selection criteria. The translated items can be viewed in Table 1.

Table 1: Items of the Strategy Questionnaire (translated from German)

no.	item	answer range (0 – 5)
1.1	We went through the tasks in sequence	not true at all – completely true
1.2	We explicitly tried to go through all tasks	not true at all – completely true
1.3	We purposefully selected tasks to work on	not true at all – completely true
1.4	If so: What criteria did you use?	open answer (optional)
1.5	We changed our strategy during learning	not true at all – completely true
1.6	If so: How?	open answer (optional)
1.7	We did not follow a specific strategy	not true at all – completely true
2.1	How much did you consider your and your learning partners' confidence in your answers?	not at all – very much
2.2	How much did you consider your and your learning partners' agreement on the answers?	not at all – very much
2.3	Are there other criteria you used to select tasks?	yes – no (binary: 0 – 1)
2.4	If yes: What were those?	open answer (optional)

Treatment

The treatment variation consisted of two factors: the provision of cognitive GA information in the form of individual assumptions about the correct answers to the learning tasks in t1 (cGAI+: yes vs. cGAI-: no) and the provision of metacognitive GA information in the form of individual confidence ratings with regard to the learning tasks in t1 (mGAI+: yes vs. mGAI-: no) during collaboration. This left us with a 2 x 2 between subjects design. Assumptions and confidence ratings were assessed within all groups during the initial assessment (learning tasks t1) and provided during collaboration to some groups, depending on experimental condition. If provided, the information of both learners was provided in separate columns labelled A and B (consistent with the side of the multi-touch table top they were standing on) next to the task in an easily comparable fashion (cf. Figure 3). By integrating the GA information into the collaborative task presented on a shared workspace, we ensured learners were able to detect relevant information and monitoring the information was associated with low additional costs (cf. Buder, 2011). Cognitive and/or metacognitive GA information provided could be changed during collaboration while an item was discussed (cf. Collaborative learning environment section). The groups provided with cognitive GA information on assumptions received information on the answers the learners had previously given in the learning tasks, and the groups provided with metacognitive GA information on confidence received information on the confidence ratings given with each answer. Cognitive GA information was spatially coded and a colored field either in the top ("true") or bottom ("false") row attached to each statement indicated the answer the learners had previously given in the learning tasks (cf. Figure 3). Metacognitive GA information was color-coded. Confident answers were indicated by a fully green field and non-confident answers were indicated by a hatched white-green field (cf. Figure 3). Figure 3 outlines how information was presented for all four research conditions.

Figure 3: Treatment conditions during collaboration (top left mGAI-/cGAI-; top right mGAI+/cGAI-; bottom left mGAI-/cGAI+; bottom right mGAI+/cGAI+)

Collaborative learning environment

During collaboration, learners interacted face to face on the multi-touch table top computer (cf. Figure 4). Within CSCL, face to face settings are of specific interest as they are common when learners jointly learn together, e.g., in preparation for exams or within schools or university courses, and may combine advantages of unmediated collaboration with the merits of computer support. Multi-touch table tops are designed for co-located learning and have great potential for collaboration (cf. Dillenbourg & Evans, 2011). By supporting face to face learning, table tops allow for multiple modes of communication including talk, gesture or action while still allowing for the benefits of computer-supported learning like interactive learning environments, embedded additional information and a shared interactive workspace. Additionally, such settings support behavioral GA, as the learners can observe each other’s activities during collaboration. While the specific design of interactive multi-touch soft- and hardware may vary considerably (cf. Higgins, Mercier, Burd, & Hatch, 2011), the technology usually allows learners to interact intuitively with the system while collaboratively working on a shared screen.

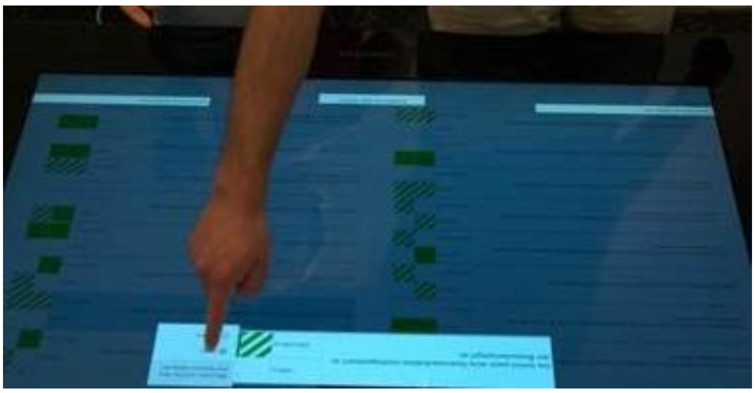
Figure 4: Two learners interacting on the multi-touch table top computer (experimental condition mGAI+/cGAI+; task visible)

During the collaboration in our study, learners were free to use the functions explained below and to discuss the material (Figure 5 shows a statechart depicting the states within the learning environment). A countdown on the screen showed the remaining time in the collaboration phase starting with 16 minutes. After the time was up, the program shut down automatically. The home screen on the table top computer showed the 16 learning tasks presented in two columns. The answers and/or confidence ratings of the learners were provided to the learners according to experimental condition. The learners had the opportunity to select specific tasks for discussion one at a time. If selected, the respective task was enlarged and shown on the top of the screen and the rest of the tasks were masked by an overlay window (TaskVisible). In this mode, the learners were able to change their previously chosen answers and/or confidence ratings (if provided) on this particular task by tapping on the respective field of the enlarged task (cf. Figure 4). Such changes were saved automatically and applied to the home screen visualization. The learners were also each able to access their original learning texts and scroll through the content, which was presented on their respective side on the table top on request (TextVisible). When doing so, they were each able to select a paragraph and send it to the middle of the screen for their partner's benefit (ParagraphVisible). When they closed the selected task, all open texts and paragraphs disappeared and were inaccessible until another task was selected (HomeScreen).

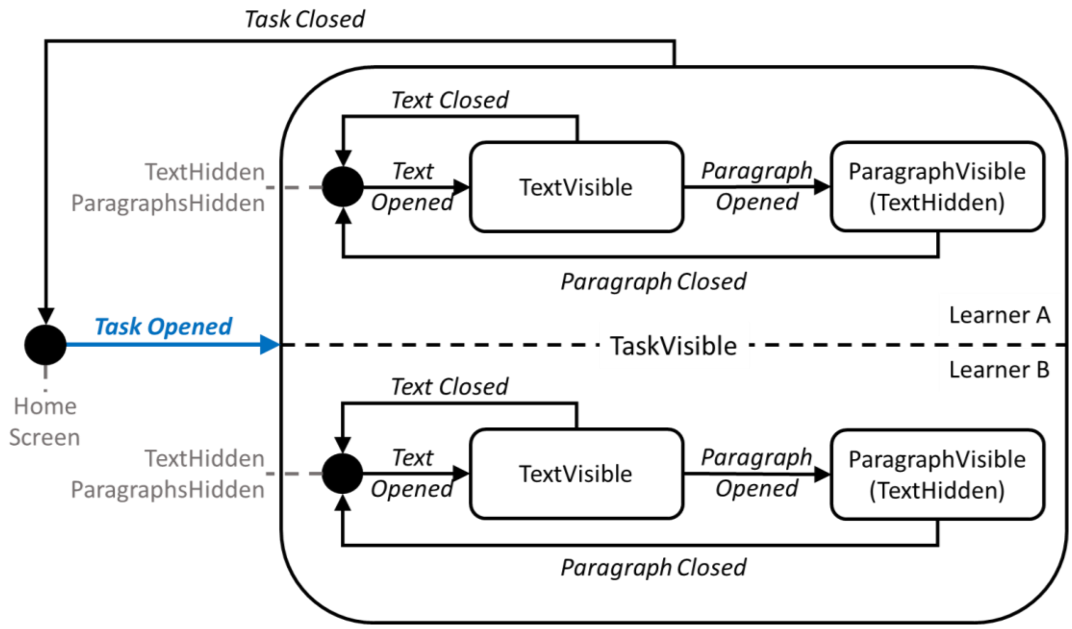

Figure 5: Statechart specifying states (rounded rectangles) and state transitions (arrows) triggered by learner-initiated events (italics) in the collaborative learning environment; central dependent variable bold-faced (blue)

Logging of collaboration data

During the collaboration on the multi-touch table top, we logged changes on the user interface. The user interface comprised multiple states (representing displayed objects), which were connected by transitions initiated by touch events conducted by the learners (cf. Figure 5). Additionally, some touch events were implemented that did not trigger state transitions (i.e., changing answers of tasks). All pre-defined touch events were logged in the form of an event log. Each event log entry consisted of an event-ID (serial number within the session), an event-type (e.g., TaskOpened, AnswerChanged), a timestamp, and a user ID (partnerA, partnerB, or dyad). For events that initiated state transitions, we additionally logged preceding and succeeding states (e.g., event: TaskOpened, preceding state: HomeScreen, succeeding state: TaskVisible; cf. Figure 5) to ease the reconstruction of state transitions and to detect inconsistencies within the logs (i.e., impossible transitions like opening a task while another is still open). Video data was used to clear up such rare issues and determine the impact on the collaboration. Where applicable, event log entries were accompanied by specifying information like task-ID or paragraph-ID. For changes in answers (assumption or confidence), the log also contained two values representing the answers before and after the change. With this log, it was possible to extract information on the learning process. Of specific interest was the selection of tasks to discuss (TaskOpened, task-ID). These could easily be matched to the previously given answers of the learners, which were saved in a specific start-log that contained information on the answer patterns at the beginning of the collaboration (i.e., the full answers and confidence ratings learners had provided regardless of information visualized).

Dependent variables: learning process

We assessed various information about the learning process and how learners proceeded during collaboration. We were especially interested in what tasks learners selected to discuss and upon what they based their discussion (i.e., conflicts/conflicting assumptions, uncertainties/uncertain assumptions). For this purpose, we used the log data (i.e., TaskOpened) from the collaboration to assess the actual tasks selected as well as questionnaire data to assess the underlying strategy as perceived by the learners.

As a first means of describing the learning process, we assessed how many tasks learners discussed during collaboration. Thus, we counted the unique selections of items to be discussed. A task was selected for discussion when learners marked the item by tapping on it (TaskOpened). It was then enlarged and the learners were able to change answers or access their initial learning texts (cf. Collaborative learning environment section). We only counted unique selections (as specified by the task-ID of each TaskOpened event) meaning that every task was counted maximally once and thus the count ranged from a (hypothetical) zero to sixteen.

We further assessed two types of regulation to assess if learners used their initial confidence and/or conflicting assumptions to regulate their selection of tasks: metacognitive (or confidence-based) regulation and conflict-based regulation. Metacognitive (confidence-based) regulation is a measure often assessed as within-subject correlation between item selection and metacognitive judgement (Thiede, 1999; Thiede et al., 2003) in metacognition research. Thus, we used such a measure to describe confidence-based regulation. We assessed item selection by looking at which items learners tapped on to indicate they would like to discuss them (event: TaskOpened); repeated selections were ignored (unique task-IDs). Tasks were coded as uncertain if at least one learner had been uncertain about their answer in t1 as logged in the start-log (independent of experimental condition). Changes in the certainty ratings during collaboration were not considered, since they always followed item selection and we only included initial selections; thus, at the time learners initially selected items, the data from t1 was the most recent data available. Due to the binarity of the data, we used φ-correlations between certainty and selection as proposed by Schraw and colleagues (Schraw, 2009; Schraw, Kuch, & Gutierrez, 2013) as measures for confidence-based regulation.

As a second measure, we assessed conflict-based regulation. Conflict-based regulation assumes that when there is a conflict, there is a collective need for clarification. Thus, conflict-based regulation was computed in a similar manner to confidence-based regulation: as a within-dyad φ-correlation between conflict-status of items (conflicting assumptions versus congruent assumptions) and item selection.

Since conflict- and confidence-based regulation coefficients both are based on correlations, learners who selected all items (or none) or did not show variation in their items with regard to confidence or conflict status had to be excluded from these analyses.

To investigate if confidence and conflict within items interact in having an impact on item selection, we assessed at what percentage the conflict-confidence combinations were selected by dividing the number of selections of each category by their occurrence. This left us with four values per dyad (2 (conflict) x 2 (confidence)) describing the percentage the four categories were selected for discussion. For analyses it was imperative that we had information for all four categories per dyad. Thus, dyads with missing categories (no occurrence) had to be excluded from these analyses.

Self-report on processes: self-developed questionnaire

We were also interested in the perspective of the learners and whether they are aware of the strategies they implement during learning and their selection processes. Thus, we used the data obtained in our new questionnaire to assess this. Factor analyses on the first part of the questionnaire (excluding open answer questions 4 and 6) revealed a two-factor solution explaining 60.83% of the variance. The items 1, 2, and 3 (negative) loaded on the first factor (38.43% of variance) and the items 5 (negative) and 7 on the other factor (explaining the remaining 22.40%). Cronbachs Alpha (cf. Cronbach, 1951) showed acceptable reliability for the first factor (unstandardized alpha = .713, whereby alpha is a rather conservative measure of internal consistency, cf. McNeish, 2017) but failed dramatically for the second (alpha = .176). Pearson correlation confirmed the unfortunate fit ($r = .10$) for this second factor and thus we abstained from interpreting it. The first factor may be interpreted as strategic item selection (vs. habitual behavior) and we used the mean score of the three items (we recoded items 1 and 2 so that a higher score indicated a more strategic selection) for testing purposes. This resulted in a measure for strategic behavior with high values (max = 5) indicating highly selective behavior and low values (min = 0) indicating a very inclusive strategy.

For part two, we were mainly interested in the first two items asking the learners if they based their selection on confidence or conflict to describe the learners' perception of their selection processes and evaluated them separately.

Dependent variables: learning outcomes

As outcome variables we assessed confidence and knowledge resulting from the collaboration. For each of the measures, we used two different tests: the learning tasks measuring changes in confidence and knowledge from pre to post on items directly relevant in the collaboration and thus very close to the treatment, and the knowledge test administered at the end of the study to assess broader and inferential knowledge about the subject as well as the confidence in that knowledge.

To measure confidence, we counted the number of confidently solved learning tasks pre- and post-collaboration. This left us with a measure between zero and sixteen for each point in time. For some analyses, we needed to combine the numbers for two learners within a dyad by computing the mean to obtain dyadic measures. This procedure results in a severe loss in variance, however, it also eliminates the influence of inflating p-values due to adding interdependent data. We additionally computed confidence gain by calculating difference values between pre- and post-test [value_t2 – value_t1]. Theoretically ranging from -16 to 16, positive values indicate a gain in confidence while negative values indicate a loss over time. While confidence gains in the learning tasks are very specific to the collaboration and very close to the learning material used during collaboration, we additionally assessed if learners gain broader knowledge on the subject and confidence in this knowledge. Thus, we used the data obtained in the knowledge test at the end of the study and computed the mean confidence level for the tasks for each participant. Since confidence was assessed on a six-point equidistant ordered response scale, we coded the data to range from 0 to 100 to receive percentage-like values for each participant.

As a further outcome variable we assessed knowledge resulting from the collaboration. Here, we counted the number of correctly solved learning tasks pre- and post-collaboration. This left us with a measure between zero and sixteen for each point in time. Again, for some analyses, we combined the numbers for two learners within a dyad by computing the mean to obtain dyadic measures, with the abovementioned advantage and disadvantage. Performance gain was also computed analogous to confidence gain by calculating difference values between pre- and post-test [value_t2 – value_t1]. Theoretically ranging from -16 to 16, positive values indicate a gain in performance while negative values indicate a loss over time. Again, while performance gains in the learning tasks are very specific to the collaboration and very close to the learning material, we additionally assessed if learners gain broader knowledge on the subject. For this, we used the data obtained in the knowledge test at the end of the study and computed the percentage of tasks solved correctly in this test for each participant.

Results

General remarks

We used different analyses to address different aspects of our research questions. Due to the partially dyadic design (decisions to discuss material were made as a dyad, while tests were taken individually) and specifics of the data (normal distribution, local interdependence, etc.), we did not conduct analyses integrating all process and outcome variables in one design. Thus, we analyzed the impact of experimental conditions on learning outcomes mediated by confidence and conflict-based regulation on a dyadic level (integrating data from both dyad members for the learning outcomes). Additionally, we tested the direct effects of the two factors (metacognitive and cognitive GA) on learning outcomes on individual as well as group level. The direct effects of the treatment on learning processes gained from log data were analyzed exclusively on dyadic level (inherently dyadic data), and data gained from the questionnaire was analyzed on the individual level. To account for distorting effects that the lack of normal distribution may have on the result, we used 10000 percentile bootstrapping in some analyses, e.g., the mediation analyses (cf. Hayes & Preacher, 2014). While other robust methods based on trimming means may be used to account for the lack of normal distribution (cf. Wilcox, 2012), current research has shown that especially in cases with highly skewed distributions (which was the case in some of our data), bootstrapping is the preferred option (Field & Wilcox, 2017). Thus, we used percentile bootstrap confidence intervals (based on 10000 bootstrap samples) as they are less susceptible to type 1 errors with smaller sample sizes than bias-corrected CIs (Fritz, Taylor, & MacKinnon, 2012) and less susceptible to outliers (Creedon & Hayes, 2015). Unless otherwise stated, alpha level was set at 5% and two-tailed tests were conducted to allow for detecting potential detrimental effects as well. Confidence intervals for effect sizes were calculated with the MBESS R-Package (Kelley, 2017).

Overall data on learning behavior

To test if providing GA information affects the raw number of discussed tasks, we counted how many of the 16 tasks were marked as discussed by each dyad and conducted a two-factorial ANOVA. Since normal distribution could not be assumed, we used 10000 percentile bootstrapping to account for the lack of normality in the data. Results show no significant effect of providing cognitive GA information on assumptions (F(1, 126) = 0.50, p = .479, $\eta_p^2 < .01$, CI 95% [0, .05]) and no significant interaction between providing cognitive GA information on assumptions and metacognitive GA information on confidence (F(1, 126) = 0.81, p = .370, $\eta_p^2 < .01$, CI 95% [0, .06]). However, we found a significant main effect of providing metacognitive GA information on confidence (F(1, 126) = 7.85, p = .006, $\eta_p^2 = .06$, CI 95% [.01, .15]). Descriptive data (Table 2) show that dyads with metacognitive GA information on confidence provided (mGAI+) tended to discuss less tasks than dyads without such information (mGAI-) with a mean difference of 1.20 tasks (CI 95% [0.35, 2.03]).

Table 2: Number of tasks selected by conditions

	number of tasks selected	
	M	*SD*
mGAI+/cGAI+	12.38	2.45
mGAI+/cGAI-	12.29	2.69
mGAI-/cGAI+	13.19	2.16
mGAI-/cGAI-	13.88	2.39

Mediation model: impact of experimental conditions on learning outcomes mediated by regulation

To address our hypotheses with regard to confidence-based regulation (H1a) and conflict-based regulation (H1b) and their effects on learning outcomes (performance gain: H2a, H2b; confidence gain: H3a, RQ1), we first computed two multiple mediation models with multi-categorical predictors predicting performance gain and confidence gain from experimental conditions mediated by confidence-based and conflict-based regulation to assess the influence of the provision of information on learning gain via regulation by using the PROCESS macro for SPSS (cf. Hayes, 2013). We used simple indicator coding for the experimental condition with the control condition (mGAI-/cGAI-: no GA information provided) as the reference. The results are depicted in Figures 6 and 7. We used percentile bootstrap confidence intervals (based on 10000 bootstrap samples) and heteroscedasticity-consistent standard errors (HC3) as described in Hayes and Cai (2007). The two models cannot be considered fully independent, since performance and confidence as outcome measures are theoretically connected, even though the specific measures did not correlate in our study (r = .018, p = .835, CI_r [-.156, .184]). Thus, we used a Bonferroni adjustment by setting the alpha-level to 2.5 percent (.025).

We first tested the model predicting confidence gain by experimental condition mediated by confidence- and conflict-based regulation (cf. Figure 6). The full model including the mediators (regulation coefficients) for confidence was statistically significant (F(5, 95) = 4.83, p < .001, R^2 = .18, CI 95% [.05, .30]), while the model without the mediators showed only a small and statistically insignificant effect, F(3, 97) = 0.68, p = .567, R^2 = .02, CI 95% [0, .08]. Further analysis revealed that learners receiving metacognitive GA information showed an increase in confidence gain mediated by confidence-based regulation ($a_{11}b_1$ = 1.21, CI 95 % [0.53, 2.07]; $a_{31}b_1$ = 0.92, CI 95% [0.33, 1.70]) (the percentile bootstrap confidence interval for the indirect effect based on 10000 bootstrap samples was entirely above zero). However, this seems to be somewhat outweighed by direct negative effects of the treatment on confidence for learners with metacognitive GA information provided (c'_1 = -1.18 (0.64), p = .069, CI 95% = [-2.46, 0.09]; c'_3 = -1.01 (0.63), p = .112, CI 95% [-2.26, 0.24]). Although these effects do not reach statistical significance separately, it is worth mentioning that there might be adverse effects and the omnibus test confirms general direct effects ($F(3, 95) = 3.74$, $p = .014$, $R^2 = 0.08$, CI 95% [.05, .21]). As expected, providing cognitive GA information led to conflict-based regulation ($a_{22} = .44$ (.09), $p < .001$, CI 95% [.26, .61]; $a_{32} = .46$ (.09), $p < .001$, CI 95% [.29, .64]), however, conflict-based regulation did not affect confidence gain ($b_2 = -1.10$ (0.81), $p = .178$, CI 95% [-2.72, 0.51]). Full data is available in appendix A.

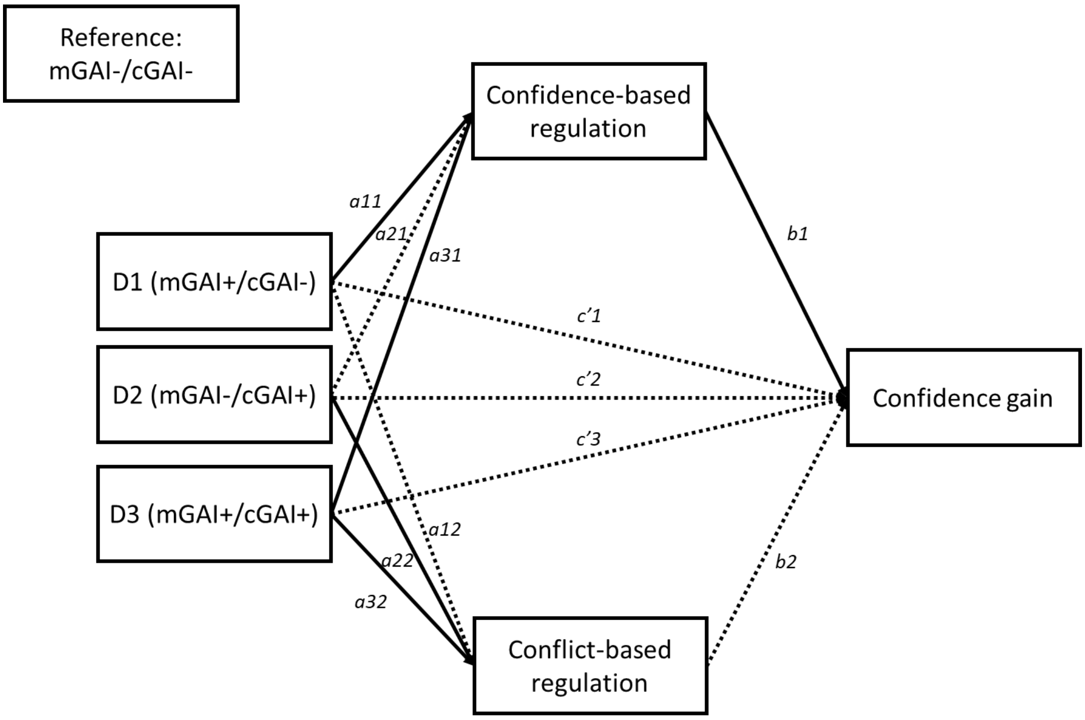

Figure 6: Confidence gain predicted by a multiple mediation model with multi-categorical predictors (significant paths are depicted with solid lines, dotted lines are not statistically significant)

For performance (cf. Figure 7), the full model including the mediators was not statistically significant ($F(5, 95) = 2.00$, $p = .086$, $R^2 = .09$, CI 95% [0, .18]), while the model without just reached statistical significance (even including the quite conservative Bonferroni correction due to two related models being tested); $F(3, 97) = 3.26$, $p = .0247$, $R^2 = .08$, CI 95% [.00, .20]). However, only one condition in the total model (excluding mediators) showed an overall significant effect: the group with only metacognitive GA information performed worse compared to the reference condition ($c_1 = -0.91$, $p = .026$, CI 95% [-1.72, -0.11]), while the other groups did not show a similar effect. Direct effects did not reach statistical significance for any of the conditions; indirect effects were marginal and insignificant (see full data in appendix B).

In the next sections we assess the impact of the treatment on regulation again in more detail – considering the two-factorial design unaccounted for in the mediation analyses.

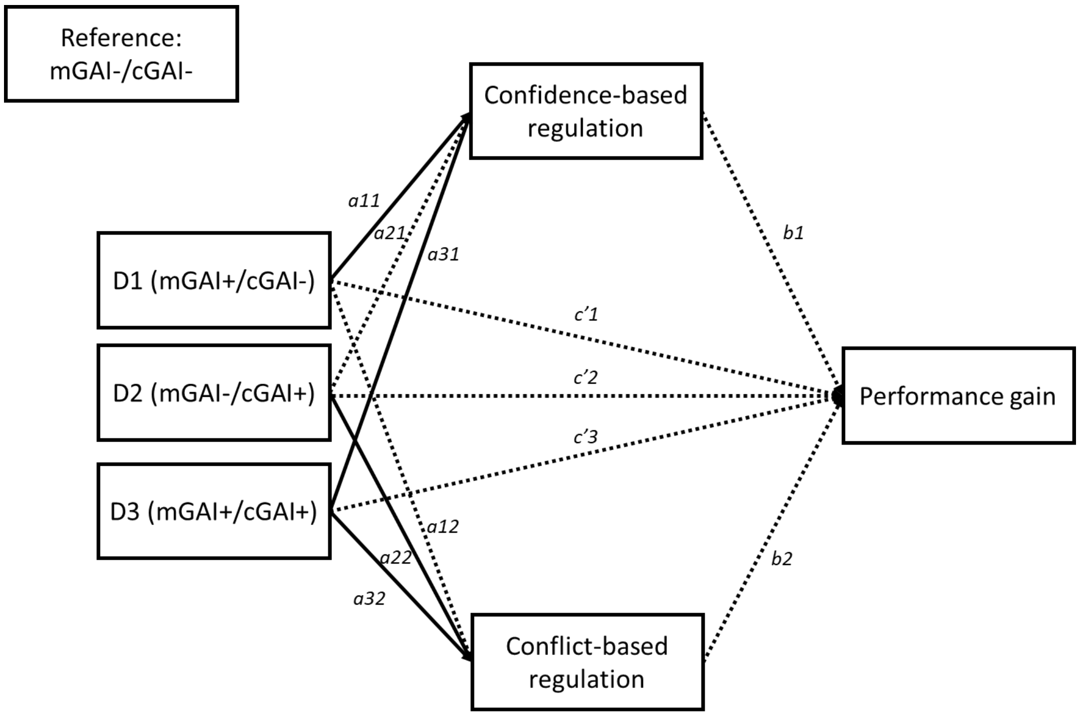

Figure 7: Performance gain predicted by a multiple mediation model with multi-categorical predictors (significant paths are depicted with solid lines, dotted lines are not statistically significant)

Learning processes 1: effects of factors on regulation

Since the mediation model could not account for the two individual factors of the experimental design (provision of metacognitive GA information on confidence and provision of cognitive GA information on assumptions), we assessed the effect of the factors on confidence-based and conflict-based regulation again separately (H1a, H1b) to additionally account for possible interaction effects (exploratory analyses of interaction). To ease the interpretation of the regulation coefficients, they were coded so that high positive coefficients meant primarily selecting uncertain or conflicting items for discussion whereas negative coefficients would mean the primary selection of certain or non-conflicting items. Values close to zero meant that no differentiation in selection was undertaken between certain and uncertain or conflicting and non-conflicting items. We conducted a two-factorial between subject MANOVA to test for the effects of the provision of metacognitive GA information and cognitive GA information on both regulation coefficients and, importantly, to test for possible interactions between the factors. It confirmed an overall effect of metacognitive GA information ($F(2, 96) = 17.68$, $p < .001$, $\eta_p^2 = .27$, CI 95% [.12, .39]) and of cognitive GA information ($F(2, 96) = 33.91$, $p < .001$, $\eta_p^2 = .41$, CI 95% [.26, .52]), but no interaction ($F(2, 96) = 1.43$, $p = .245$, $\eta_p^2 = .03$, CI 95% [0, .11]). Metacognitive GA information only had an effect on confidence-based regulation ($F(1, 97) = 35.73$, $p < .001$, $\eta_p^2 = .27$, CI 95% [.13, .40]) but not on conflict-based regulation ($F(1, 97) = 0.31$, $p = .582$, $\eta_p^2 < .01$, CI 95% [0, .06]) and cognitive GA information had an effect solely on conflict-based regulation ($F(1, 97) = 67.67$, $p < .001$, $\eta_p^2 = .41$, CI 95% [.26, .53]), but not on confidence-based regulation ($F(1, 97) = 0.03$, $p = .870$, $\eta_p^2 <.001$, CI 95% [0, .03]). The interaction of the two factors did not have an effect on confidence-based regulation ($F(1, 97) < .01$, $p = .964$, $\eta_p^2 < .001$, CI 95% [.00, .00]) or on conflict-based regulation ($F(1, 97) = 2.87$, $p = .093$, $\eta_p^2 = .03$, CI 95% [0, .12]). Because the data was not normally distributed and thus not fully suited for the analysis used, we confirmed the results using the MANOVA.RM R-Package, bootstrapping the data with a parametric bootstrapping approach (Friedrich, Konietschke, & Pauly, 2017) and 10000 bootstrapping iterations. The results confirm the multivariate effect of metacognitive GA information ($F_{MATS}(2, 96) = 37.14$, $p < .001$) and cognitive GA information ($F_{MATS}(2, 96) = 63.45$, $p < .001$), as well as the lack of interaction between the two factors ($F_{MATS}(2, 96) = 2.96$, $p = .241$).

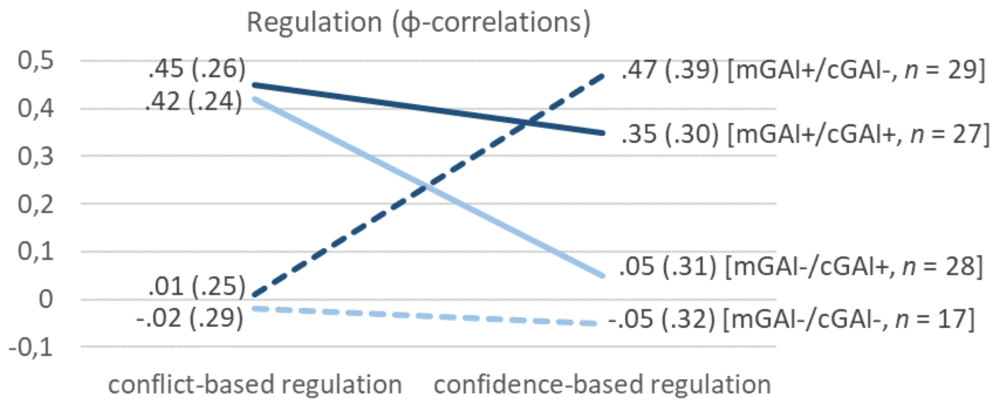

Figure 8: Means (standard deviations) of the regulation coefficients for conflict-based regulation and confidence-based regulation for all groups. Solid lines are for cGAI+ conditions (dotted lines for cGAI-), darker shade lines are for mGAI+ conditions (lighter shade for mGAI-)

Descriptive statistics show that with both types of GA information available (mGAI+/cGAI+), learners use both regulation types and use them fairly equally, while without GA information (mGAI-/cGAI-) there is no correlation between either confidence or conflict with discussion (cf. Figure 8). It is worth mentioning that since the regulation coefficients were based on correlations, dyads who selected all tasks for discussion had to be excluded from the analyses (hypothetically this would have also been true for learners with no variation in terms of certainty or conflict, however, this did not occur). This was the case especially in the group without GA information as support (mGAI-/cGAI-; cf. Figure 8).

Learning processes 2: integration of metacognitive and cognitive information per experimental condition

While we used regulation coefficients to look at how learners base their decisions on confidence levels and conflict status separately, we were further interested in whether they integrated the information (exploratory analyses of interaction). Unfortunately, the full design becomes even more inflated when including these two within-subject factors (confidence level and conflict status) in addition to the two between subject factors (metacognitive GA information on confidence and cognitive GA information on assumptions). Thus, we analyzed the information separately for each experimental condition. To further include the information and especially the interaction of conflict status and confidence on affecting discussion, we mapped the discussion rate (%) for confidence x conflict-status for all groups separately (cf. Figure 9). Since dyads lacking any of the four examined patterns (e.g., conflicting assumptions with both learners being confident) had to be excluded from the analyses, the *n* within the groups may deviate from the data described above.

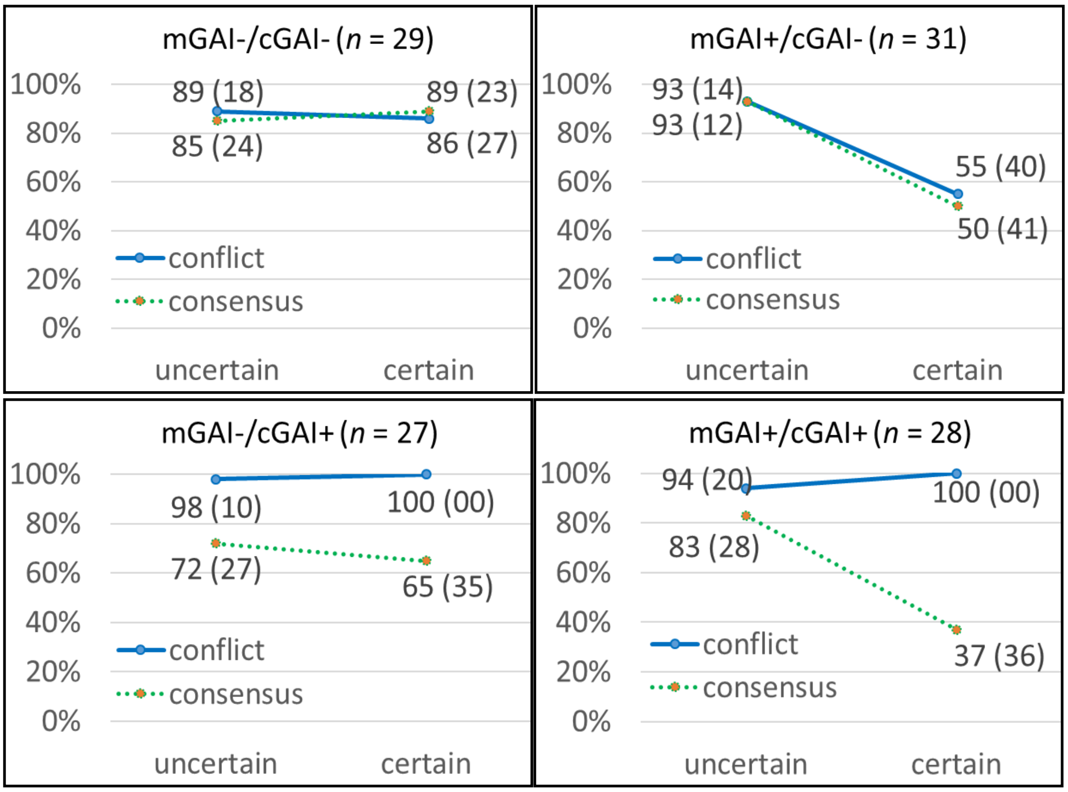

Figure 9: Mean of percentage of tasks discussed (standard deviations) by confidence and conflict status per condition

The results show that for the condition without GA information (mGAI-/cGAI-), neither confidence ($F(1, 28) < .01$, $p = .951$, $\eta_p^2 < .01$, CI 95% [.00, .04]) nor conflict status ($F(1, 28) < .01$, $p = .994$, $\eta_p^2 < .01$, CI 95% [.00, .00]), nor their interaction ($F(1, 28) = 1.24$, $p = .275$, $\eta_p^2 = .04$, CI 95% [0, .24]) had an effect on the rate the items were discussed, which was rather high in general (cf. Figure 9, top left). The condition with only metacognitive GA information on confidence (mGAI+/cGAI-) showed a significant effect of confidence on the discussion rate ($F(1, 30) = 35.43$, $p < .001$, $\eta_p^2 = .54$, CI 95% [.27, .69]), but neither of conflict status ($F(1, 30) = 0.64$, $p = .430$, $\eta_p^2 = .02$, CI 95% [0, .19]) nor an interaction ($F(1, 30) = 0.61$, $p = .443$, $\eta_p^2 = .02$, CI 95% [0, .19]) (cf. Figure 9, top right). The condition with only cognitive GA information on assumptions (mGAI-/cGAI+) showed a significant effect of conflict status on discussion rate ($F(1, 26) = 38.51$, $p < .001$, $\eta_p^2 = .60$, CI 95% [.32, .73]), but not of confidence ($F(1, 26) = 0.36$, $p = .555$, $\eta_p^2 = .01$, CI 95% [0, .19]) nor an interaction effect ($F(1, 26) = 1.21$, $p = .281$, $\eta_p^2 = .05$, CI 95% [0, .25]) (cf. Figure 9, bottom left). For the group with both types of GA information visible (mGAI+/cGAI+), we found significant main effects for both confidence ($F(1, 27) = 26.41$, $p < .001$, $\eta_p^2 = .49$, CI 95% [.20, .66]) and conflict status ($F(1, 27) = 46.69$, $p < .001$, $\eta_p^2 = .63$, CI 95% [.37, .76]) and an interaction effect ($F(1, 27) = 43.65$, $p < .001$, $\eta_p^2 = .62$, CI 95% [.35, .74]) (cf. Figure 9, bottom right). Viewing the descriptive data in Figure 9, we can see that while conflict might have a universal effect on discussion percentage (albeit marginalized if accounting for/extracting the interaction), the main effect of confidence level can well be explained by the interaction. Since the data was heavily skewed for some cells (mainly ceiling effects), the results need to be interpreted with caution and merely provide an indicator for the effect the patterns have on topic discussions within each experimental condition.

Learning processes 3: strategic procedure from questionnaire data

We further inquired if learners perceived their own behavior as being strategic and selective (H1a, H1b, self-report). The data showed relatively low scores on the scale ranging from more habitual learning behavior on the low end (e.g., working through tasks in sequential order, not choosing specific tasks, but working through all of them) to more selective behavior on the high end (cf. Table 3). However, statistical analyses revealed differences between the groups. A two-factorial ANOVA showed no effect for metacognitive GA information on confidence ($F(1, 256) = 1.30$, $p = .255$, $\eta_p^2 < .01$, CI 95% [0, .04]) or the interaction of the two types of GA information ($F(1, 256) = 1.52$, $p = .219$, $\eta_p^2 < .01$, CI 95% [0, .04]), but a statistically significant effect of cognitive GA information on assumption ($F(1, 256) = 16.23$, $p < .001$, $\eta_p^2 = .06$, CI 95% [.02, .12]). For the latter effect, mean distances were roughly 0.70 (CI 95% [0.36, 1.04]) with the conditions with cognitive GA information on assumptions (cGAI+) reporting higher levels of selective behavior. We used 10000 percentile bootstrapping samples for estimation to account for the lack of normal distribution in the data. Descriptive statistics can be found in Table 3.

Table 3: Descriptive statistics on questionnaire data

	selective behavior scale	selection due to confidence	selection due to conflict
	M (SD)	*M (SD)*	*M (SD)*
mGAI+/cGAI+	2.04 (1.56)	3.03 (1.31)	3.89 (1.24)
mGAI+/cGAI-	1.55 (1.45)	2.94 (1.41)	3.66 (1.28)
mGAI-/cGAI+	2.05 (1.31)	2.95 (0.98)	4.03 (0.93)
mGAI-/cGAI-	1.14 (1.31)	3.19 (0.92)	3.75 (1.20)

To assess if learners perceived their procedure to be based on confidence, we conducted a 2-way ANOVA (10000 percentile bootstrapping). It showed no differences between the factor levels for either the provision of metacognitive GA information on confidence ($F(1, 256) = 0.33$, $p = .566$, $\eta_p^2 < .01$, CI 95% [0, .02]), the provision of cognitive GA information on assumptions ($F(1, 256 = 0.23$, $p = .623$, $\eta_p^2 < .01$, CI 95% [0, .02]) or the interaction of both ($F(1, 256) = 1.23$, $p = .269$, $\eta_p^2 < .01$, CI 95% [0, .04]). We used the same procedure to assess whether learners perceived their procedure to be based on conflict. Again, we found no significant effects of providing metacognitive GA information on confidence ($F(1, 256) = 0.62$, $p = .431$, $\eta_p^2 < .01$, CI 95% [0, .03]), cognitive GA information on assumptions ($F(1, 256) = 3.10$, $p = .080$, $\eta_p^2 = .01$, CI 95% [0, .05]), or their interaction ($F(1, 256) = 0.03$, $p = .857$, $\eta_p^2 < .01$, CI 95% [0, .01]). Descriptive statistics for both items are shown in Table 3.

Learning outcomes: effects of factors on cognitive and metacognitive learning outcomes

Due to the complexity of the design, we used dyadic performance gain and confidence gain in the learning tasks in the mediation model, thereby losing variance between measurements (post – pre) and dyad members (Mean[learner A, learner B]). Since the mediation models could not provide full details on the variance of direct effects, we additionally analyzed the data again without mediators (testing H2a, H2b, H3a, and RQ1), thereby including measurement points (within-subject factor time: pre, post) and learners separately in a 2x2x2 MANOVA with repeated measures on one factor (cf. Table 4). Since normal distribution could not be assumed for some of the cells in the design, the results need to be interpreted with caution.

Results showed a highly significant multivariate effect of time (*F*(2, 255) = 204.70, *p* < .001, η_p^2 = .62, CI95% [.54, .67]), which can be seen in univariate analyses to be due to both gains in confidence (*F*(1, 255) = 112.81, *p* < .001, η_p^2 = .31, CI 95% [.22, .39]) and performance (*F*(1, 255) = 321.08, *p* < .001, η_p^2 = .56, CI 95% [.48, .62]). However, as shown in Table 4, no other main or interaction effects were statistically significant.

Table 4: Multivariate effects of time and the two types of GA information on learning outcomes (performance and confidence) [2x2x2 MANOVA]

factors (type)	*F*(2, 255)	*p*	η_p^2	CI 95%
main effects				
time (within)	204.70	< .001	.62	.54, .67
metacognitive GA information (between)	2.00	.138	.02	.00, .05
cognitive GA information (between)	0.03	.967	< .01	.00, .001
first order interactions				
metacognitive * cognitive GA information	0.46	.632	< .01	.00, .03
time * metacognitive GA information	1.95	.144	.02	.00, .05
time * cognitive GA information	1.30	.273	.01	.00, .04
second order interaction				
time * metacognitive * cognitive GA information	2.40	.093	.02	.00, .06

Due to partial interdependence of the dyad members (cf. Table 5) as measured with the intra-class correlation (ICC; Shrout & Fleiss, 1979), we confirmed the results on the dyad level (discarding inner-dyadic variance). The data can be viewed in Appendix C. The results are consistent with the individual level analyses with only an effect of time of measurement (pre or post collaboration) being statistically significant.

Interestingly, intra-class correlation coefficients differed between conditions: While as expected, performance and confidence level prior to the collaboration were not interdependent within dyads (ICC values ranged from -.15 to .19 for performance and .04 to .24 for confidence and were statistically non-significant), and overall showed statistically significant interdependences post-collaboration, the post-collaboration values differed between the conditions (cf. Table 5). While there is no straightforward interpretation, it seems that learners with metacognitive GA information (mGAI+) are more strongly interdependent when it comes to metacognitive outcomes (confidence level post), than learners without such information (mGAI-). However, with regard to cognitive outcomes, the opposite seems to be the case: Learners without metacognitive GA information (mGAI-) seem to be more interdependent than learners with metacognitive GA information (mGAI+). However, this seems to depend on cognitive information as well. Further, it is interesting, that in one condition only (mGAI+/cGAI+), partners' performance scores seem not to be interrelated. With regard to confidence, cognitive GA information does not seem to play a major role as the ICC values are quite similar between groups with (cGAI+) and without (cGAI-).

Table 5: ICC values (and *p*-values) [*n*] for the learning tasks post collaboration

	performance post			confidence post		
	cGAI+ ICC (*p*)	cGAI- ICC (*p*)	overall ICC (*p*)	cGAI+ ICC (*p*)	cGAI- ICC (*p*)	overall ICC (*p*)
mGAI+	-.205 (.875) [*n* = 32]	.358 (.017) [*n* = 34]	.126 (.154) [*n* = 66]	.458 (.003) [*n* = 32]	.610 (<.001) [*n* = 34]	.537 (<.001) [*n* = 66]
mGAI-	.580 (<.001) [*n* = 32]	.499 (.001) [*n* = 32]	.534 (<.001) [*n* = 64]	.365 (.017) [*n* = 32]	.241 (.086) [*n* = 32]	.307 (.006) [*n* = 64]
overall	.246 (.023) [*n* = 64]	.444 (<.001) [*n* = 66]	.352 (<.001) [*n* = 130]	.392 (.001) [*n* = 64]	.417 (<.001) [*n* = 66]	.400 (<.001) [*n* = 130]

We conducted another two-factorial MANOVA testing the impact of GA information on post-learning confidence (H3a, RQ1) and performance levels (H2a, H2b) assessed in the knowledge test at the end of the study. We conducted the analysis with individual data since intra-class-correlation measures showed no indication of interdependence between learners (values were entirely below .20 and non-significant). The test for normality showed no significant deviation from normality, after one outlier was excluded from the calculation. Results showed that neither the provision of metacognitive GA information on confidence ($F(2, 254) = 0.23$, $p = .796$, $\eta_p^2 < .01$, CI 95% [0, .02]), nor the provision of cognitive GA information on assumptions ($F(2, 254) = 1.91$, $p = .150$, $\eta_p^2 = .02$, CI 95% [0, .04]), nor their interaction ($F(2, 254) = 1.61$, $p = .203$, $\eta_p^2 = .01$, CI 95% [0, .04]) had an impact on the results in the post test. Descriptive data can be viewed in Table 6.

Table 6: descriptive data on confidence and performance in knowledge test (percentage)

		mean performance knowledge test (0 - 100)		mean confidence knowledge test (0 - 100)	
experimental condition	*n*	*M*	*SD*	*M*	*SD*
mGAI+/cGAI+	64	45.23	5.43	63.05	14.69
mGAI+/cGAI-	67	47.00	6.30	63.93	14.04
mGAI-/cGAI+	64	46.28	6.64	59.86	13.33
mGAI-/cGAI-	64	46.22	5.21	64.95	14.26

Discussion

Results of the study

In our study we compared two types of knowledge-related group awareness (GA) information within a CSCL scenario regarding their impact on learning processes and outcomes: cognitive content information enabling conflict identification and metacognitive confidence information enabling identification of perceived lacks of knowledge. Overall, we found that the type of information provided guided the selection of material to discuss. Learners used the information provided to make study decisions (confirming H1a & H1b) and even somewhat seemed to integrate the information if both were provided. This ties in with GA research and shows the power of GA information to guide learning decisions. Interestingly, the interaction effect between conflict and confidence on the discussion rates if both types of information are provided hints at a universal effect of conflict, but confidence guiding learning decisions primarily when there is consensus amongst learners (or when there is no information on conflict provided). Because the data base for these results is quite scarce, follow-up studies should revisit these effects and look more closely at how both types of information are perceived and processed by the learners.

While selection processes were adapted as expected, this did not have the expected impact on learning. Overall, learners gained knowledge while working collaboratively within the learning environment, but that seemed to not depend on the GA information provided or regulation processes (rejecting H2a & H2b). While previous studies showed a similar lack of effect of GA tools on learning outcomes (e.g., Buder & Bodemer, 2008; Engelmann & Hesse, 2011; Schnaubert & Bodemer, 2016), it is still surprising that even regarding the precise items they worked with, discussing conflicts and uncertainties did not have beneficial effects on performance, even if research on metacognition and socio-cognitive conflicts would suggest differently (cf. Cognitive and metacognitive information in collaboration section).

However, looking into the discussion rates, we can see that even if unsupported learners do not focus on conflicts and uncertainties, they still look into almost all the items. Thus, the benefit in providing GA information may be an issue of efficiency rather than effectiveness (especially the provision of metacognitive information seems to reduce the number of items discussed). In our study, all learners had the same amount of time and were not allowed to terminate the process prematurely. Because we kept the learning time stable to avoid time as an influential factor, there might have been no urgent need to structure and decide on a strategy and some learners may have thought that using a strategy like looking at all items would suffice. Tighter time constraints should force learners to make strategic decisions. Without it, one potentially beneficial effect may have been lost, because learners were able to attend to all material. More realistically, learners could have been given the opportunity to terminate the session prematurely. This would be more realistic, since in practice, there are usually some time constraints students face when working collaboratively on a task, but they usually still have some control over the time and effort they put into learning. While effects could then not be attributed to the quality of collaboration, but maybe merely the quantity, this procedure would be in line with research on self-regulated learning, which assumes that deciding when to terminate study is an important learning decision (e.g., Metcalfe & Kornell, 2005).

Another explanation for the lack of effect on learning gain would be that learners may have tried to resolve conflicts by quick consensus building activities (Weinberger & Fischer, 2006) and may have wanted to complete the task rather than gaining knowledge. Such an approach is less beneficial and may have been supported by the experimental setup (participation was rewarded, but there may have been no internal value in gaining knowledge). Additionally, providing information on differences of assumptions has been shown to foster this approach under some circumstances (Gijlers et al., 2009) and this might have hampered the potentially beneficial effects of attending to conflicts in the conditions with cognitive GA information available.

Providing cognitive GA information on assumptions did not affect post-learning confidence levels (RQ1), but we did find that metacognitive GA information had an indirect impact on confidence mediated by confidence-based regulation, however, this was partially masked by a negative (albeit not statistically significant) direct effect on this measure for both groups with metacognitive GA information available (thus only partially confirming H3a). Consequently, it seems possible that if learners do not strategically use the information on own uncertainties to steer their learning process, having it saliently visible throughout learning may foster the preservation of these uncertainties. While such effects would have to be backed up by further research (especially in light of the still scarce evidence), we have to consider the possibility of detrimental effects of GA information if it is not used strategically.

Such detrimental effects of providing additional information may also be explained by limitations of the working memory (e.g., Sweller, 1994), especially with novice learners as was the case in our study (cf. Kalyuga, 2013). If information is provided that does not serve the learning process, this may put an additional strain on the cognitive system and thus hamper germane learning processes (Sweller, van Merrienboer, & Paas, 1998). This may even be more relevant within CSCL scenarios, where the collaboration may add to the complexity of the situation and transactive activities produce additional costs (cf. Dillenbourg & Bétrancourt, 2006; Kirschner, Sweller, Kirschner, & Zambrano R., 2018). In addition, a shared workspace forces learners to negotiate not only their understanding of a task, but also negotiate and coordinate their interaction with the learning environment (cf. Dourish & Bellotti, 1992). However, mental overload may only partially explain the found effects, since knowledge gain was not affected, but rather the metacognitive evaluation. Thus, it may well be that providing GA information on uncertainties may lead to further uncertainties because learners become aware of their gaps in knowledge and if they do not use this (meta-level) knowledge to clear up uncertainties, they may be unsettled. To avoid such effects, providing metacognitive GA information on confidence may be accompanied by more direct instructions to ensure their usage to benefit learning (e.g., by prompting or scripting, cf. Kollar, Wecker, & Fischer, 2018).

Interestingly, while we found clear patterns about how learners used GA information to structure their learning process in the log files, their self-reports did not differ in terms of usage of one type of information or the other. There are various possible explanations for this. The easiest would be to assume that learners are just not aware of the strategies they use or are not able to reproduce them after learning (flaws on self-report, e.g., Nisbett & Wilson, 1977). However, it is equally possible that if not provided, learners reconstruct the information on certainty and assumptions during collaboration and that this repetition of the process results in different assumptions and certainties than before. We do know that metacognitive judgments may change over time and with attention to material (e.g., Koriat & Levy-Sadot, 2001; Vernon & Usher, 2003), and especially assumptions made under high levels of uncertainty (like guessing) seem likely to change merely due to chance.

Additionally, information gathered during the collaboration may affect the learners' certainties and assumptions, e.g., conflicting cognitive information from the learning partners may make learners uncertain. The learners may then use the adjusted information as a basis for regulatory processes. If such changes are only initiated during collaboration, they will be unaccounted for in the experimental log data initially assessed and may lead us to false assumptions about the regulatory processes taking place. Especially if learners do not reproduce but newly construct the information, it seems logical that they take information gathered during collaboration into account (e.g., provided answers of their learning partner may affect confidence; information gathered during collaboration may affect assumptions). Thus, the learners in our study may well have been using this newly constructed information very precisely, but without a chance for us to account for these cognitive and metacognitive changes when relying on pre-collaboration data. That said, reconstructing the information on the fly may well produce a more accurate picture of the learners' knowledge than using information produced prior to collaboration (although merely reproducing it could also be essentially flawed, e.g., due to constrains in memory). However, reconstructing information takes up valuable resources. Metacognitive monitoring processes and the need to keep the information mentally present while performing other cognitive activities essentially for learning may overstrain learners (Valcke, 2002). Although prior studies with similar material did not show an overall rise in mental effort if metacognitive information was not visualized (Schnaubert & Bodemer, 2017), it is possible that in a collaborative scenario the additional effort of maintaining metacognitive and cognitive information of both partners may tip the balance and produce a load on the cognitive system that was too high to handle properly.

While the log data paints a different picture than the self-reports, there is another aspect which may have had a potential impact on these results. While providing GA information set focus on conflicting issues and uncertainties, the assessment of the GA information alone may have had similar effects. Individual research has shown repeatedly that merely assessing metacognitive ratings (like confidence ratings or judgments of learning) affects memory (Soderstrom, Clark, Halamish, & Bjork, 2015) and study choices (Mitchum, Kelley, & Fox, 2016; Schnaubert & Bodemer, 2017) by prompting the assessed monitoring processes in the first place. While this somewhat hampers the external validity of the study results with regard to the effects of providing GA information, assessing and transforming relevant information is a crucial part of GA tools (Bodemer et al., 2018; Buder & Bodemer, 2008). Using potentially reactive direct assessment methods of the target concepts as opposed to non-obtrusively attained data (like analyzing available products of student work, e.g., essays: M. Erkens et al., 2016) may also yield benefits as transformation processes may be kept to a minimum. Transformation processes inevitably integrate external information or algorithms and thus are a source of external feedback for the learners, especially if they are complex like common in the field of learning analytics. This has several implications: First, such processes require external information of high quality like expert models or computational algorithms adjusted to the content domain. This makes such tools inflexible and somewhat impractical for various settings (including school settings) as it hampers transferability to different learning materials and domains. Second, forgoing transformation makes the whole process of obtaining and providing GA information instantly transparent and easily acceptable for learners as they are in control of the information provided. Thus, providing untampered information about the learners' take on their knowledge can easily be interpreted by target learners, as they provided the information in the first place (we acknowledge that this may differ between learners and may thus sometimes be less straightforward when learners interpret their learning partner's information). This focus on information provided by the learners themselves rather than on externally provided information further strengthens the notion of guiding without governing, which is focal for self-regulation and agency and sets GA apart from other instructional methods like feedback and more explicit guidance.

Limitations of the study

There are some limitations to the study that need to be addressed. Possibly the most obvious one concerns the regulation measure. Although the effects on regulation seem strong and stable, it is worth mentioning that the individual regulation coefficients might be somewhat error prone. The correlation coefficients were based on 16 observations each – some of them unevenly distributed on either of the variables, which is a rather low number of observations; to get more stable measures, more data points would be needed. However, 16 items was already a large set and we wanted to avoid overloading learners, which could pressure them into using the provided information as guidance due to the sheer amount of information. Interestingly, most of the dyads we had to exclude from the calculations because they discussed all tasks (and no correlation could be computed) were in the group without GA information support. We excluded them from the calculation, but it is fair to point out that, technically, they did not differentiate between conflicting and non-conflicting items or certain and uncertain items, since they did look at all tasks without fault. This further supports the notion that guidance occurred when GA information support was provided.

We have some further limitations to the data that are worth mentioning. Some of the dependent variables assessed showed highly skewed distributions (e.g., ceiling effects) and although bootstrapping was used whenever possible, this does not account for all problems associated with this.

Due to the dyadic and complex design, we also abstained from using multilevel analyses, even though our subjects worked together in dyads and thus have shown interdependences in some instances. The learning processes we investigated are not affected, since they are dyadic by nature, but learning gains and confidence levels in the learning tasks were affected. ICC values showed that learners aligned their confidence levels especially when information on confidence was provided and less so if it wasn't. However, the opposite seemed to be the case for performance. This is in line with previous findings, where metacognitive information led to higher interdependencies with regard to confidence levels and lower interdependencies with regard to performance (Schnaubert & Bodemer, 2018). However, in our study, one condition especially seemed to break ranks with regard to performance: learners provided with both types of information were not more similar within than between dyads. One interpretation is that while learners may align their assumptions especially if cognitive information is provided, information on confidence might allow learners to maintain differing assumptions by reducing socio-cognitive conflict due to salience of low confidence (confidence in assumptions is believed to be a factor in perceiving socio-cognitive conflict and thus the need to align opinions; cf. Lee & Kwon, 2001; Lee et al., 2003). However, descriptively looking into ICC values may help with interpreting the results but cannot provide definitive conclusions without inspecting collaboration processes. To account for the fact that we violated the independency assumption of the statistical tests used, we additionally performed the analyses on a dyadic level. While this does come with other problems (cf. Janssen, Erkens, Kirschner, & Kanselaar, 2011), it does account for underestimated *p*-values due to interrelated data. Overall, while we might have overestimated effects on learning outcomes by ignoring the hierarchical structure and using individual data, we found no effects to speak of on learning outcomes anyway, rendering the critique merely academic.

Another decision that needs to be discussed is the usage of binary information with regard to the information portrayed in our GA tool. Within metacognition research, it is common to assess metacognitive information on a more fine-grained scale and binary confident scales are assumed to be inferior for research purposes (cf. Dinsmore & Parkinson, 2013). However, in this kind of research, the aim usually is to assess metacognitive information to study it, not to guide learning processes. As argued above, within GA tools, the information has to be presented in an easy to understand way and aligned to the intended guidance mechanisms (cf. Bodemer, 2011). Theoretically, it would have been possible to assess the information on a more fine-grained scale (for research purposes) but present it in a binary fashion (for the learners' benefit). However, any transformation chosen may not align with what learners individually perceive as relevant differences between certainty and uncertainty. Confronted with a binary choice, learners are able to define a threshold of what is "certain enough", which aligns with the binary decision of what needs no further attention. While this seems fitting for supporting task selection, it may hamper more fine-grained decision making like prioritizing (cf. Schnaubert & Bodemer, 2017).

Ultimately, tool design needs to strike a balance between portraying rich information and keeping processing costs low. In terms of guidance, such design decisions are crucial and need to be scrutinized in detail, because how data is assessed, transformed and visualized may have a huge impact on guidance mechanisms (for an example see the work on representational guidance mechanisms, e.g., Suthers, 2001; Suthers & Hundhausen, 2003). While our research focused on the information portrayed, this cannot be fully detached from design decisions made that may suggest certain learning behaviors more than others. Further research should explicitly target such decisions to find the best possible balance between richness and usability.

Implications for practitioners and CSCL design

The results of this research suggest that different types of GA information lead to very different approaches to learning material. Co-located learners working in learning groups may profit from information about each other by being supported by GA tools to detect conflicting information or perceived uncertainties. Integrating the GA information into the collaborative task at hand proved useful as this ensured that the information did not divert attention from the task (cf. Buder, 2011) and our results showed that learners used the information for content selection. Further, learners were able to integrate both types of information when provided. This suggests that educators wanting learners to focus on different types of information simultaneously may use tools providing different types of information without overburdening the learners – at least in cases where the information is of low complexity, easy to understand and integrated in the learning environment. Scaling up the information by including richer data still needs to be done with caution, as more complex data may exponentially increase the mental effort needed to integrate various kinds of information and thus warrants further research to explicitly test the boundaries for specific learning arrangements.

Looking at the interaction effect between conflict status and confidence with regard to the discussion rates, the data cautiously suggests that especially confident learners may be engaged in collaborative learning processes by including information about conflicts and it may even be worth deliberately pairing them with partners disagreeing to encourage engagement with the learning content (an example of a script containing such a paring algorithm with relation to opinions is the argue graph script, cf. Jermann & Dillenbourg, 2003).

While the guidance effects were clearly visible in the study, our results did not suggest benefits of GA information with regard to learning outcomes. One of the strengths of GA support to guide learning is that it builds on self-regulatory skills and gives learners the freedom to adapt the learning process to their needs. However, this may just also be its greatest weakness: learners with low self-regulatory skills or learners struggling with other cognitive or collaborative processes involved (like argumentation) may not be able to profit from such support in terms of learning gains. There is ongoing debate about how much and directive support to grant learners (e.g., Kirschner & van Merriënboer, 2013). In our study, learners had support in regulating their study choices, but no support in conducting the studying itself. While the setup and tool provided opportunities for collaborative activities such as joining attention towards aspects by pointing or referring to a shared external representation or sharing information from each individual text, no explicit structure for discussing the content was provided. Although more explicit guidance is often seen in stark contrast to implicit guidance mechanisms such as awareness tools, this view is not unchallenged (see the debate in Wise & Schwarz, 2017). By carefully looking into guidance effects as well as learning outcomes, our research points towards an integrative approach: while the GA information portrayed by the GA tool in our study managed to draw the learners' attention towards relevant aspects of the learning material and thus supported the metacognitive activity of selecting relevant content and allocating study resources, the process of collaboratively dealing with the material may need more support. Scaffolding beneficial collaborative activities (e.g., by scripting, Kollar et al., 2018) to complement GA support seems to be especially fitting, as scripts may be used to specify activities and support their execution (cf. Weinberger, Ertl, Fischer, & Mandl, 2005), while the GA information helped learners to identify relevant conditions within the collaborative situation and focused their attention on specific content. If collaborative learning is supported by implicit and explicit guidance mechanisms in unison, both types of support not only need to match the goals of the educator and the cognitive processes involved, but also need to be aligned to each other. For instance, cognitive information on assumptions may warrant support in argumentation and conflict resolution processes (e.g., Stegmann, Wecker, Weinberger, & Fischer, 2012; Stegmann, Weinberger, & Fischer, 2007), while metacognitive information on confidence may be accompanied by support in explaining content or asking questions (e.g., King, 1992; Rosenshine, Meister, & Chapman, 1996).

The use of these guidance mechanisms requires educators to be aware of their learners' individual characteristics as these may determine if and how learners benefit from the support provided. Unfortunately, research on individual characteristics and their relation to implicit guidance approaches are rare (for a notable exception see Heimbuch & Bodemer, 2018). Looking more thoroughly into learners' skill sets and other potentially relevant characteristics may be needed to better adapt the support towards the learners' needs.

Closing remark

Our study focused on guidance effects of GA information within a shared learning environment and thus on how learners structure their learning with the aid of GA information by interacting with the learning material within the learning environment rather than what cognitive and collaborative processes happen during collaboration. We are aware that this perspective does not directly target cognitive and metacognitive processes relevant to learning, but merely their behavioral outputs. From a self-regulation perspective, guiding study behavior and choosing what aspects of learning material are worth allocating study time to are crucial aspects of knowledge acquisition and have thus received a lot of attention in metacognition research. Within research on collaborative learning, the focus often lies in communication and interaction processes between learners as they are the core of collaboration and additionally allow us to study cognitive processes in more detail as they are often externalized for the learning partners' benefit. While this was beyond the scope of the research presented in this study, looking more deeply into the processes of sharing and building knowledge may help explain the lack of effect on knowledge and help us to better understand how learners use cognitive and metacognitive GA information for learning. Finally, our study provided relevant insight into the effects of different types of GA information portrayed within knowledge-related GA tools. However, more studies are needed that systematically vary specific aspects of these tools in different settings to advance our knowledge on mechanisms and boundary conditions of GA tools to improve their effectiveness in supporting CSCL.

Acknowledgements We would like to thank Christian Schlusche, M.Sc., for the extensive technical support he provided.

Appendices

Appendix A

```
Run MATRIX procedure:
************* PROCESS Procedure for SPSS Release 2.16.1 ***********
          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2013). www.guilford.com/p/hayes3
**************************************************************************
Model = 4
    Y = Confiden [annot.: confidence gain learning tasks]
    X = GA_condi [annot.: group awareness condition]
   M1 = reg_conf [annot.: conflict-based regulation]
   M2 = reg_rc_u [annot.: confidence-based regulation]
Sample size: 101
Coding of categorical X variable for analysis:
 GA_condi       D1       D2       D3
     1.00      .00      .00      .00 [annot.: no visual.]
     2.00     1.00      .00      .00 [annot.: confidence visual.]
     3.00      .00     1.00      .00 [annot.: assumption visual.]
     4.00      .00      .00     1.00 [annot.: both visualizations]
**************************************************************************
Outcome: reg_conf
Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .6455      .4166      .0672    21.9210     3.0000    97.0000      .0000
Model
              coeff         se          t          p       LLCI       ULCI
Constant     -.0180      .0736     -.2443      .8075     -.1641      .1281
D1            .0316      .0875      .3610      .7189     -.1422      .2054
D2            .4370      .0869     5.0272      .0000      .2645      .6096
D3            .4639      .0900     5.1544      .0000      .2853      .6425
**************************************************************************
Outcome: reg_rc_u
Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .5314      .2824      .1118    11.8293     3.0000    97.0000      .0000
Model
              coeff         se          t          p       LLCI       ULCI
Constant     -.0495      .0795     -.6229      .5348     -.2072      .1082
D1            .5227      .1086     4.8130      .0000      .3071      .7382
D2            .1043      .0990     1.0533      .2948     -.0922      .3008
D3            .3960      .0992     3.9911      .0001      .1991      .5930
**************************************************************************
Outcome: Confiden
Model Summary
          R       R-sq        MSE          F        df1        df2          p
      .4190      .1756     3.4924     4.8263     5.0000    95.0000      .0006
Model
```

	coeff	se	t	p	LLCI	ULCI
Constant	3.0357	.4631	6.5547	.0000	2.1163	3.9552
reg_conf	-1.1034	.8127	-1.3578	.1777	-2.7168	.5099
reg_rc_u	2.3108	.6090	3.7947	.0003	1.1019	3.5198
D1	-1.1831	.6437	-1.8378	.0692	-2.4610	.0949
D2	.3715	.6236	.5957	.5528	-.8665	1.6096
D3	-1.0112	.6296	-1.6061	.1116	-2.2610	.2387

************************** TOTAL EFFECT MODEL ******************

Outcome: Confiden

Model Summary

R	R-sq	MSE	F	df1	df2	p
.1476	.0218	4.0584	.6783	3.0000	97.0000	.5674

Model

	coeff	se	t	p	LLCI	ULCI
Constant	2.9412	.4305	6.8322	.0000	2.0868	3.7956
D1	-.0101	.5788	-.0175	.9861	-1.1589	1.1386
D2	.1303	.5937	.2194	.8268	-1.0481	1.3086
D3	-.6078	.5862	-1.0370	.3023	-1.7712	.5556

***************** TOTAL, DIRECT, AND INDIRECT EFFECTS **********

Relative total effects of X of Y

	coeff	se	t	p	LLCI	ULCI
D1	-.0101	.5788	-.0175	.9861	-1.1589	1.1386
D2	.1303	.5937	.2194	.8268	-1.0481	1.3086
D3	-.6078	.5862	-1.0370	.3023	-1.7712	.5556

Omnibus test of total effect of X on Y

R-sq	F	df1	df2	p
.0218	.6783	3.0000	97.0000	.5674

=====

Relative direct effects of X on Y

	coeff	se	t	p	LLCI	ULCI
D1	-1.1831	.6437	-1.8378	.0692	-2.4610	.0949
D2	.3715	.6236	.5957	.5528	-.8665	1.6096
D3	-1.0112	.6296	-1.6061	.1116	-2.2610	.2387

Omnibus test of direct effect of X on Y

R-sq	F	df1	df2	p
.0829	3.7386	3.0000	95.0000	.0137

=====

Relative indirect effect(s) of X on Y through: reg_conf

	Effect	SE(boot)	LLCI	ULCI
D1	-.0349	.1210	-.3022	.2076
D2	-.4822	.3609	-1.1770	.2520
D3	-.5118	.3994	-1.3219	.2619
Omnibus	-.4398	.3530	-1.1730	.2576

Relative indirect effect(s) of X on Y through: reg_rc_u

	Effect	SE(boot)	LLCI	ULCI
D1	1.2078	.3968	.5276	2.0702
D2	.2410	.2394	-.2098	.7508
D3	.9152	.3496	.3341	1.6971
Omnibus	.6012	.2533	.2386	1.2128

******************** ANALYSIS NOTES AND WARNINGS *************

Number of bootstrap samples for percentile bootstrap confidence intervals: 10000
Level of confidence for all confidence intervals in output: 95.00
NOTE: Some cases were deleted due to missing data. The number of such cases was: 29
NOTE: All standard errors for continuous outcome models are based on the HC3 estimator
------ END MATRIX -----

Appendix B

Run MATRIX procedure:
************* PROCESS Procedure for SPSS Release 2.16.1 ***********
Written by Andrew F. Hayes, Ph.D. www.afhayes.com
Documentation available in Hayes (2013). www.guilford.com/p/hayes3
****
Model = 4
Y = Performa [annot.: performance gain learning tasks]
X = GA_condi [annot.: group awareness condition]
M1 = reg_conf [annot.: conflict-based regulation]
M2 = reg_rc_u [annot.: confidence-based regulation]
Sample size: 101
Coding of categorical X variable for analysis:

GA_condi	D1	D2	D3
1.00	.00	.00	.00 [annot.: no visual.]
2.00	1.00	.00	.00 [annot.: confidence visual.]
3.00	.00	1.00	.00 [annot.: assumption visual.]
4.00	.00	.00	1.00 [annot.: both visualizations]

****
Outcome: reg_conf
Model Summary

R	R-sq	MSE	F	df1	df2	p
.6455	.4166	.0672	21.9210	3.0000	97.0000	.0000

Model

	coeff	se	t	p	LLCI	ULCI
Constant	-.0180	.0736	-.2443	.8075	-.1641	.1281
D1	.0316	.0875	.3610	.7189	-.1422	.2054
D2	.4370	.0869	5.0272	.0000	.2645	.6096
D3	.4639	.0900	5.1544	.0000	.2853	.6425

****
Outcome: reg_rc_u
Model Summary

R	R-sq	MSE	F	df1	df2	p
.5314	.2824	.1118	11.8293	3.0000	97.0000	.0000

Model

	coeff	se	t	p	LLCI	ULCI
Constant	-.0495	.0795	-.6229	.5348	-.2072	.1082
D1	.5227	.1086	4.8130	.0000	.3071	.7382
D2	.1043	.0990	1.0533	.2948	-.0922	.3008
D3	.3960	.0992	3.9911	.0001	.1991	.5930

****
Outcome: Performa

Model Summary

R	R-sq	MSE	F	df1	df2	p
.2929	.0858	1.9170	2.0006	5.0000	95.0000	.0855

Model

	coeff	se	t	p	LLCI	ULCI
Constant	1.4880	.3323	4.4774	.0000	.8283	2.1478
reg_conf	-.2775	.4309	-.6439	.5212	-1.1330	.5780
reg_rc_u	-.1408	.3846	-.3661	.7151	-.9043	.6227
D1	-.8314	.4326	-1.9219	.0576	-1.6903	.0274
D2	.0645	.4649	.1388	.8899	-.8584	.9875
D3	.1845	.4101	.4498	.6539	-.6297	.9987

************************** TOTAL EFFECT MODEL ******************

Outcome: Performa

Model Summary

R	R-sq	MSE	F	df1	df2	p
.2862	.0819	1.8855	3.2637	3.0000	97.0000	.0247

Model

	coeff	se	t	p	LLCI	ULCI
Constant	1.5000	.3337	4.4956	.0000	.8378	2.1622
D1	-.9138	.4053	-2.2544	.0264	-1.7183	-.1093
D2	-.0714	.4536	-.1575	.8752	-.9718	.8289
D3	.0000	.4200	.0000	1.0000	-.8336	.8336

***************** TOTAL, DIRECT, AND INDIRECT EFFECTS **********

Relative total effects of X of Y

	coeff	se	t	p	LLCI	ULCI
D1	-.9138	.4053	-2.2544	.0264	-1.7183	-.1093
D2	-.0714	.4536	-.1575	.8752	-.9718	.8289
D3	.0000	.4200	.0000	1.0000	-.8336	.8336

Omnibus test of total effect of X on Y

R-sq	F	df1	df2	p
.0819	3.2637	3.0000	97.0000	.0247

=====

Relative direct effects of X on Y

	coeff	se	t	p	LLCI	ULCI
D1	-.8314	.4326	-1.9219	.0576	-1.6903	.0274
D2	.0645	.4649	.1388	.8899	-.8584	.9875
D3	.1845	.4101	.4498	.6539	-.6297	.9987

Omnibus test of direct effect of X on Y

R-sq	F	df1	df2	p
.0600	2.5732	3.0000	95.0000	.0586

=====

Relative indirect effect(s) of X on Y through: reg_conf

	Effect	SE(boot)	LLCI	ULCI
D1	-.0088	.0466	-.1346	.0609
D2	-.1213	.1981	-.5348	.2491
D3	-.1287	.2116	-.5691	.2606
Omnibus	-.1106	.1920	-.5087	.2550

Relative indirect effect(s) of X on Y through: reg_rc_u

	Effect	SE(boot)	LLCI	ULCI
D1	-.0736	.2066	-.4875	.3385
D2	-.0147	.0552	-.1387	.0998

D3 -.0558 .1556 -.3731 .2495
Omnibus -.0366 .1148 -.2841 .1832

******************** ANALYSIS NOTES AND WARNINGS *************
Number of bootstrap samples for percentile bootstrap confidence intervals: 10000
Level of confidence for all confidence intervals in output: 95.00
NOTE: Some cases were deleted due to missing data. The number of such cases was: 29
NOTE: All standard errors for continuous outcome models are based on the HC3 estimator
------ END MATRIX -----

Appendix C

factors (type)	$F(2, 125)$	p	η_p^2
main effects			
time (within)	179.70	< .001	.74
metacognitive GA information (between)	1.57	.211	.03
cognitive GA information (between)	0.03	.974	< .01
first order interactions			
metacognitive * cognitive GA information	0.36	.700	< .01
time * metacognitive GA information	1.75	.177	.03
time * cognitive GA information	1.15	.320	.02
second order interaction			
time * metacognitive * cognitive GA information	1.95	.147	.03

References

Apache Software Foundation. (2013). *Apache CouchDB* [Computer Program]. Wakefield, MA: Apache Software Foundation.

Ariel, R., Dunlosky, J., & Bailey, H. (2009). Agenda-based regulation of study-time allocation: When agendas override item-based monitoring. *Journal of Experimental Psychology: General*, *138*(3), 432–447. https://doi.org/10.1037/a0015928

Bodemer, D. (2011). Tacit guidance for collaborative multimedia learning. *Computers in Human Behavior*, *27*(3), 1079–1086. https://doi.org/10.1016/j.chb.2010.05.016

Bodemer, D., & Dehler, J. (2011). Group awareness in CSCL environments. *Computers in Human Behavior*, *27*(3), 1043–1045. https://doi.org/10.1016/j.chb.2010.07.014

Bodemer, D., Janssen, J., & Schnaubert, L. (2018). Group awareness tools for computer-supported collaborative learning. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (pp. 351–358). New York, NY: Routledge/Taylor & Francis.

Bodemer, D., & Scholvien, A. (2014). Providing knowledge-related partner information in collaborative multimedia learning: Isolating the core of cognitive group awareness tools. In C.-C. Liu, H. Ogata, S. C. Kong, & A. Kashihara (Eds.), *Proceedings of the 22nd International Conference on Computers in Education ICCE 2014* (pp. 171–179). Nara, JP.

Buchs, C., Butera, F., Mugny, G., & Darnon, C. (2004). Conflict elaboration and cognitive outcomes. *Theory Into Practice*, *43*(1), 23–30. https://doi.org/10.1207/s15430421tip4301_4

Buder, J. (2011). Group awareness tools for learning: Current and future directions. *Computers in Human Behavior*, *27*(3), 1114–1117. https://doi.org/10.1016/j.chb.2010.07.012

Buder, J., & Bodemer, D. (2008). Supporting controversial CSCL discussions with augmented group awareness tools. *International Journal of Computer-Supported Collaborative Learning*, *3*(2), 123–139. https://doi.org/10.1007/s11412-008-9037-5

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, *65*(3), 245–281.

Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology. Learning, Memory & Cognition*, *27*(6), 1491. https://doi.org/10.1037/0278-7393.27.6.1491

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, *1*(1), 69–84. https://doi.org/10.1007/s11409-006-6894-z

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: American Psychological Association.

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in Psychology*, *9*(C), 287–299. https://doi.org/10.1016/S0166-4115(09)60059-5

Crano, W., & Prislin, R. (2006). Attitudes and persuasion. *Annual Review of Psychology*, *57*, 345–374. https://doi.org/10.1146/annurev.psych.57.102904.190034

Creedon, P. J., & Hayes, A. F. (2015, May). *Small sample mediation analysis: How far can you push the bootstrap?* Presented at the Annual conference of the Association for Psychological Science, New York, NY, US.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Dehler, J., Bodemer, D., & Buder, J. (2007). Fostering audience design of computer-mediated knowledge communication by knowledge mirroring. In C. Chinn, G. Erkens, & S. Puntambekar (Eds.), *Proceedings of the 7th Computer Supported Collaborative Learning Conference* (pp. 168–170). New Brunswick, Canada: International Society of the Learning Sciences.

Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2009). Providing group knowledge awareness in computer-supported collaborative learning: Insights into learning mechanisms. *Research and Practice in Technology Enhanced Learning*, *4*(2), 111–132. https://doi.org/10.1142/S1793206809000660

Dehler, J., Bodemer, D., Buder, J., & Hesse, F. W. (2011). Guiding knowledge communication in CSCL via group knowledge awareness. *Computers in Human Behavior*, *27*(3), 1068–1078. https://doi.org/10.1016/j.chb.2010.05.018

Dillenbourg, P., & Bétrancourt, M. (2006). Collaboration load. In J. Elen & R. E. Clark (Eds.), *Handling complexity in learning environments: Research and theory* (pp. 142–163). Amsterdam, NL: Elsevier.

Dillenbourg, P., & Evans, M. (2011). Interactive tabletops in education. *International Journal of Computer-Supported Collaborative Learning*, *6*(4), 491–514. https://doi.org/10.1007/s11412-011-9127-7

Dillenbourg, P., & Hong, F. (2008). The mechanics of CSCL macro scripts. *International Journal of Computer-Supported Collaborative Learning*, *3*(1), 5–23. https://doi.org/10.1007/s11412-007-9033-1

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-enhanced learning* (pp. 3–19). Dordrecht, NL: Springer.

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review*, *20*(4), 391–409. https://doi.org/10.1007/s10648-008-9083-6

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

Doise, W., & Mugny, G. (1984). Sociocognitive conflict. In M. Argyle (Ed.), *The Social Development of the Intellect* (Vol. 10, pp. 77–101). Amsterdam: Pergamon. Retrieved from https://www.sciencedirect.com/science/article/pii/B978008030215750010X

Doise, W., Mugny, G., & Perret-Clermont, A.-N. (1975). Social interaction and the development of cognitive operations. *European Journal of Social Psychology*, *5*(3), 367–383. https://doi.org/10.1002/ejsp.2420050309

Dourish, P., & Bellotti, V. (1992). Awareness and coordination in shared workspaces. In M. Mantel & R. Baecker (Eds.), *Proceedings of the 1992 ACM conference on Computer-supported cooperative work* (pp. 107–114). Toronto, Canada: ACM Press. https://doi.org/10.1145/143457.143468

Drachsler, H., & Greller, W. (2012). The pulse of learning analytics understandings and expectations from the stakeholders. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge* (pp. 120–129). New York, NY, USA: ACM. https://doi.org/10.1145/2330601.2330634

Dunlosky, J., & Hertzog, C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. *Psychology and Aging*, *15*(3), 462–474. https://doi.org/10.1037/0882-7974.15.3.462

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Los Angeles, CA: Sage Publications.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction*, *22*(4), 271–280. https://doi.org/10.1016/j.learninstruc.2011.08.003

Efklides, A., Samara, A., & Petropoulou, M. (1999). Feeling of difficulty: An aspect of monitoring that influences control. *European Journal of Psychology of Education*, *14*(4), 461–476. https://doi.org/10.1007/BF03172973

Engelmann, T., Dehler, J., Bodemer, D., & Buder, J. (2009). Knowledge awareness in CSCL: A psychological perspective. *Computers in Human Behavior*, *25*(4), 949–960. https://doi.org/10.1016/j.chb.2009.04.004

Engelmann, T., & Hesse, F. W. (2010). How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. *International Journal of Computer-Supported Collaborative Learning*, *5*(3), 299–319. https://doi.org/10.1007/s11412-010-9089-1

Engelmann, T., & Hesse, F. W. (2011). Fostering sharing of unshared knowledge by having access to the collaborators' meta-knowledge structures. *Computers in Human Behavior*, *27*, 2078–2087. https://doi.org/10.1016/j.chb.2011.06.002

Erkens, G., Jaspers, J., Prangsma, M., & Kanselaar, G. (2005). Coordination processes in computer supported collaborative writing. *Computers in Human Behavior*, *21*(3), 463–486. https://doi.org/10.1016/j.chb.2004.10.038

Erkens, M., Bodemer, D., & Hoppe, H. U. (2016). Improving collaborative learning in the classroom: Design and evaluation of a text mining based grouping and representing. *International Journal of Computer-Supported Collaborative Learning*, *11*(4), 387–415. https://doi.org/10.1007/s11412-016-9243-5

Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, *16*(1), 88–92. https://doi.org/10.3758/PBR.16.1.88

Field, A. P., & Wilcox, R. R. (2017). Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. *Behaviour Research and Therapy*, *98*, 19–38. https://doi.org/10.1016/j.brat.2017.05.013

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist*, *48*(1), 56–66. https://doi.org/10.1080/00461520.2012.748005

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

Fransen, J., Kirschner, P. A., & Erkens, G. (2011). Mediating team effectiveness in the context of collaborative learning: The importance of team and task awareness. *Computers in Human Behavior*, *27*(3), 1103–1113. https://doi.org/10.1016/j.chb.2010.05.017

Fraundorf, S. H., & Benjamin, A. S. (2016). Conflict and metacognitive control: The mismatch-monitoring hypothesis of how others' knowledge states affect recall. *Memory*, *24*(8), 1108–1122. https://doi.org/10.1080/09658211.2015.1069853

Friedrich, S., Konietschke, F., & Pauly, M. (2017). MANOVA.RM: Analysis of multivariate data and repeated measures designs (Version 0.2.1).

Fritz, M. S., Taylor, A. B., & MacKinnon, D. P. (2012). Explanation of two anomalous results in statistical mediation analysis. *Multivariate Behavioral Research*, *47*(1), 61–87. https://doi.org/10.1080/00273171.2012.640596

Gijlers, H. (2005). *Confrontation and co-construction: Exploring and supporting collaborative scientific discovery learning with computer simulations* (Doctoral dissertation). University of Twente, Enschede. Retrieved from http://doc.utwente.nl/50896/

Gijlers, H., Saab, N., van Joolingen, W. R., de Jong, T., & van Hout-Wolters, B. H. A. M. (2009). Interaction between tool and talk: How instruction and tools support consensus building in collaborative inquiry-learning environments. *Journal of Computer Assisted Learning*, *25*(3), 252–267. https://doi.org/10.1111/j.1365-2729.2008.00302.x

Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Journal of Educational Technology & Society*, *15*(3), 42–57.

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (2009). A growing sense of "agency." In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 1–4). New York, NY: Routledge.

Hancock, T. E., Stock, W. A., & Kulhavy, R. W. (1992). Predicting feedback effects from response-certitude estimates. *Bulletin of the Psychonomic Society*, *30*(2), 173–176. https://doi.org/10.3758/BF03330431

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: The Guilford Press.
Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods*, *39*(4), 709–722. https://doi.org/10.3758/BF03192961
Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, *67*(3), 451–470. https://doi.org/10.1111/bmsp.12028
Heimbuch, S., & Bodemer, D. (2018). Interaction of guidance types and the Need for Cognitive Closure in wiki-based learning. *PeerJ*, *6:e5541*. https://doi.org/10.7717/peerj.5541
Hesse, F. (2007). Being told to do something or just being aware of something? An alternative approach to scripting in CSCL. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting computer-supported communication of knowledge-cognitive, computational and educational perspectives* (pp. 91–98). New York, NY: Springer.
Higgins, S., Mercier, E., Burd, E., & Hatch, A. (2011). Multi-touch tables and the relationship with collaborative classroom pedagogies: A synthetic review. *International Journal of Computer-Supported Collaborative Learning*, *6*(4), 515–538. https://doi.org/10.1007/s11412-011-9131-y
Hines, J. C., Touron, D. R., & Hertzog, C. (2009). Metacognitive influences on study time allocation in an associative recognition task: An analysis of adult age differences. *Psychology and Aging*, *24*(2), 462–475. https://doi.org/10.1037/a0014417
Hunt, D. P. (2003). The concept of knowledge and how to measure it. *Journal of Intellectual Capital*, *4*(1), 100–113. https://doi.org/10.1108/14691930310455414
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., & Salonen, P. (2011). Socially shared metacognition of dyads of pupils in collaborative mathematical problem-solving processes. *Learning and Instruction*, *21*(3), 379–393. https://doi.org/10.1016/j.learninstruc.2010.05.002
Janssen, J., & Bodemer, D. (2013). Coordinated computer-supported collaborative learning: Awareness and awareness tools. *Educational Psychologist*, *48*(1), 40–55. https://doi.org/10.1080/00461520.2012.749153
Janssen, J., Erkens, G., & Kanselaar, G. (2007). Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in Human Behavior*, *23*(3), 1105–1125. https://doi.org/10.1016/j.chb.2006.10.005
Janssen, J., Erkens, G., & Kirschner, P. A. (2011). Group awareness tools: It’s what you do with it that matters., *27*(3). https://doi.org/10.1016/j.chb.2010.06.002
Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2011). Multilevel analysis in CSCL research. In S. Puntambekar, G. Erkens, & C. Hmelo-Silver (Eds.), *Analyzing Interactions in CSCL* (pp. 187–205). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4419-7710-6_9
Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist*, *48*(1), 25–39. https://doi.org/10.1080/00461520.2012.748006

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. https://doi.org/10.1007/s11412-016-9238-2

Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., … Järvenoja, H. (2015). Enhancing socially shared regulation in collaborative learning groups: Designing for CSCL regulation tools. *Educational Technology Research and Development*, *63*(1), 125–142. https://doi.org/10.1007/s11423-014-9358-1

Jarvis, B. G. (2012). *MediaLab (Version 2012)* [Computer Software]. New York: Empirisoft Corporation.

Jermann, P., & Dillenbourg, P. (2003). Elaborating new arguments through a CSCL script. In J. Andriessen, M. Baker, & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments* (pp. 205–226). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-0781-7_8

Johnson, D. W., & Johnson, R. T. (2009a). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher*, *38*(5), 365–379. https://doi.org/10.3102/0013189X09339057

Johnson, D. W., & Johnson, R. T. (2009b). Energizing learning: The instructional power of conflict. *Educational Researcher, 38*(1), 37–51. https://doi.org/10.3102/0013189X08330540

Johnson, D. W., Johnson, R. T., & Tjosvold, D. (2000). Constructive controversy: The value of intellectual opposition. In M. Deutsch & P. T. Coleman (Eds.), *The handbook of conflict resolution: Theory and practice* (pp. 65–85). San Francisco, CA: Jossey-Bass.

Johnson, R., Brooker, C., Stutzman, J., Hultman, D., & Johnson, D. W. (1985). The effects of controversy, concurrence seeking, and individualistic learning on achievement and attitude change. *Journal of Research in Science Teaching, 22*(3), 197–205. https://doi.org/10.1002/tea.3660220302

Kalyuga, S. (2013). Effects of learner prior knowledge and working memory limitations on multimedia learning. *Procedia - Social and Behavioral Sciences*, *83*, 25–29. https://doi.org/10.1016/j.sbspro.2013.06.005

Kelley, K. (2017). *MBESS (Version 4.0.0).*

King, A. (1992). Facilitating elaborative learning through guided student-generated questioning. *Educational Psychologist, 27*(1), 111–126. https://doi.org/10.1207/s15326985ep2701_8

King, A. (2007). Scripting collaborative learning processes: A cognitive perspective. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting computer-supported collaboratorive Learning: Cognitive, computational and educational perspectives* (pp. 13–37). New York, NY: Springer.

Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano R., J. (2018). From Cognitive Load Theory to Collaborative Cognitive Load Theory. *International Journal of Computer-Supported Collaborative Learning*, *13*(2), 213–233. https://doi.org/10.1007/s11412-018-9277-y

Kirschner, P. A., & van Merriënboer, J. J. G. (2013). Do learners really know best? Urban legends in education. *Educational Psychologist*, *48*(3), 169–183. https://doi.org/10.1080/00461520.2013.804395

Kollar, I., Wecker, C., & Fischer, F. (2018). Scaffolding and scripting (computer-supported) collaborative learning. In F. Fischer, C. E. Hmelo-Silver, S. R.

Goldman, & P. Reimann (Eds.), *International Handbook of the Learning Sciences* (pp. 340–350). New York, NY: Routledge/Taylor & Francis.
Koriat, A. (2012). The relationships between monitoring, regulation and performance. *Learning and Instruction*, *22*(4), 296–298. https://doi.org/10.1016/j.learninstruc.2012.01.002
Koriat, A., Adiv, S., & Schwarz, N. (2015). Views that are shared with others are expressed with greater confidence and greater fluency independent of any social influence. *Personality and Social Psychology Review*, *20*(2), 176–193. https://doi.org/10.1177/1088868315585269
Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *27*(1), 34–53. https://doi.org/10.1037//0278-7393.27.1.34
Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General, 135*(1), 36–69. https://doi.org/10.1037/0096-3445.135.1.36
Kornell, N., & Metcalfe, J. (2006). Study efficacy and the region of proximal learning framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(3), 609–622. https://doi.org/10.1037/0278-7393.32.3.609
Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, *1*(4), 279–308. https://doi.org/10.1007/BF01320096
Kulhavy, R. W., Stock, W. A., Hancock, T. E., Swindell, L. K., & Hammrich, P. L. (1990). Written feedback: Response certitude and durability. *Contemporary Educational Psychology, 15*(4), 319–332. https://doi.org/10.1016/0361-476X(90)90028-Y
Leclercq, D. (1983). Confidence marking: Its use in testing. *Evaluation in Education*, *6*(2), 161–287. https://doi.org/10.1016/0191-765X(82)90011-8
Leclercq, D. (1993). Validity, reliability, and acuity of self-assessment in educational testing. In D. Leclercq & J. E. Bruno (Eds.), *Item banking: Interactive testing and self-assessment* (pp. 114–131). Berlin, Germany: Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-58033-8_11
Leclercq, D., & Poumay, M. (2004). Objective assessment of subjectivity: Degrees of certainty and partial knowledge. Presented at the 2nd Biennial Meeting of the EARLI Special Interest Group 16 Metacognition, Amsterdam, NL.
Lee, G., & Kwon, J. (2001). What do we know about students' cognitive conflict in science classroom: A theoretical model of cognitive conflict process. In P. A. Rubba, J. A. Rye, W. J. Di Biase, & B. A. Crawford (Eds.), *Proceedings of the 2001 Annual Meeting of the Association for the Education of Teachers in Science* (pp. 309–325). Costa Mesa, CA: Association for the Education of Teachers in Science.
Lee, G., Kwon, J., Park, S.-S., Kim, J.-W., Kwon, H.-G., & Park, H.-K. (2003). Development of an instrument for measuring cognitive conflict in secondary-level science classes. *Journal of Research in Science Teaching, 40*(6), 585–603. https://doi.org/10.1002/tea.10099
Levine, J. M., Resnick, L. B., & Higgins, E. T. (1993). Social foundations of cognition. *Annual Review of Psychology*, *44*(1), 585–612. https://doi.org/10.1146/annurev.ps.44.020193.003101

Lowry, N., & Johnson, D. W. (1981). Effects of controversy on epistemic curiosity, achievement, and attitudes. *The Journal of Social Psychology*, *115*(1), 31–43. https://doi.org/10.1080/00224545.1981.9711985

Maki, R. H. (1998a). Predicting performance on text: Delayed versus immediate predictions and tests. *Memory & Cognition*, *26*(5), 959–964. https://doi.org/10.3758/BF03201176

Maki, R. H. (1998b). Test predictions over text material. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 117–144). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

McNeish, D. (2017). Thanks coefficient alpha, we'll take it from here. *Psychological Methods*, *32*(3). https://doi.org/10.1037/met0000144

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*(1), 174–179. https://doi.org/10.3758/PBR.15.1.174

Metcalfe, J., & Finn, B. (2011). People's hypercorrection of high-confidence errors: Did they know it all along? *Journal of Experimental Psychology. Learning, Memory & Cognition*, *37*(2), 437–448. https://doi.org/10.1037/a0021962

Metcalfe, J., & Kornell, N. (2005). A region of proximal learning model of study time allocation. *Journal of Memory and Language*, *52*(4), 463–477. https://doi.org/10.1016/j.jml.2004.12.001

Miller, M., & Hadwin, A. (2015). Scripting and awareness tools for regulating collaborative learning: Changing the landscape of support in CSCL. *Computers in Human Behavior*, *52*, 573–588. https://doi.org/10.1016/j.chb.2015.01.050

Mitchum, A. L., Kelley, C. M., & Fox, M. C. (2016). When asking the question changes the ultimate answer: Metamemory judgments change memory. *Journal of Experimental Psychology: General*, *145*(2), 200–219. https://doi.org/10.1037/a0039923

Mugny, G., Butera, F., Sanchez-Mazas, M., & Perez, J. A. (1995). Judgments in conflict: The conflict elaboration theory of social influence. In B. Boothe, R. Hirsig, A. Helminger, & R. Volkart (Eds.), *Perception, evaluation, interpretation*. Seattle, WA: Hogrefe & Huber.

Mugny, G., & Doise, W. (1978). Socio-cognitive conflict and structure of individual and collective performances. *European Journal of Social Psychology*, *8*(2), 181–192. https://doi.org/10.1002/ejsp.2420080204

Nelson, T. O., Dunlosky, J., Graf, A., & Narens, L. (1994). Utilization of metacognitive judgments in the allocation of study during multitrial learning. *Psychological Science*, *5*(4), 207–213. https://doi.org/10.1111/j.1467-9280.1994.tb00502.x

Nelson, T. O., & Leonesio, J. R. (1988). Allocation of self-paced study time and the "labor-in-vain effect." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(4), 676–686. https://doi.org/10.1037/0278-7393.14.4.676

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *Psychology of learning & motivation* (Vol. 26, pp. 125–173). New York, NY: Academic Press.

Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. *Psychological Bulletin*, *125*(6), 737–759. https://doi.org/10.1037/0033-2909.125.6.737

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

Nova, N., Wehrle, T., Goslin, J., Bourquin, Y., & Dillenbourg, P. (2007). Collaboration in a multi-user game: impacts of an awareness tool on mutual modeling.

Multimedia Tools and Applications, *32*(2), 161–183. https://doi.org/10.1007/s11042-006-0065-8

Panadero, E., & Järvelä, S. (2015). Socially shared regulation of learning: A review. *European Psychologist*, *20*(3), 190–203. https://doi.org/10.1027/1016-9040/a000226

Price, P. C., & Stone, E. R. (2004). Intuitive evaluation of likelihood judgment producers: Evidence for a confidence heuristic. *Journal of Behavioral Decision Making*, *17*(1), 39–57. https://doi.org/10.1002/bdm.460

Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research*, *66*(2), 181–221. https://doi.org/10.3102/00346543066002181

Sangin, M., Molinari, G., Nüssli, M.-A., & Dillenbourg, P. (2008). Knowing what the peer knows: The differential effect of knowledge awareness on collaborative learning performance of asymmetric pairs. In P. Dillenbourg & M. Specht (Eds.), *Times of convergence. Technologies across learning contexts* (pp. 384–394). Berlin, Germany: Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-87605-2_43

Sangin, M., Molinari, G., Nüssli, M.-A., & Dillenbourg, P. (2011). Facilitating peer knowledge modeling: Effects of a knowledge awareness tool on collaborative learning outcomes and processes. *Computers in Human Behavior*, *27*(3), 1059–1067. https://doi.org/10.1016/j.chb.2010.05.032

Schnaubert, L., & Bodemer, D. (2016). How socio-cognitive information affects individual study decisions. In C.-K. Looi, J. Polman, U. Cress, & P. Reimann (Eds.), *Transforming Learning, Empowering Learners: The International Conference of the Learning Sciences (ICLS) 2016* (pp. 274–281). Singapore, SG: International Society of the Learning Sciences.

Schnaubert, L., & Bodemer, D. (2017). Prompting and visualising monitoring outcomes: Guiding self-regulatory processes with confidence judgments. *Learning and Instruction*, *49*, 251–262. https://doi.org/10.1016/j.learninstruc.2017.03.004

Schnaubert, L., & Bodemer, D. (2018). What interdependence can tell us about collaborative learning: A statistical and psychological perspective. *Research and Practice in Technology Enhanced Learning*, *13*(1), 1–18. https://doi.org/10.1186/s41039-018-0084-x

Scholvien, A., & Bodemer, D. (2013). Information cueing in collaborative multimedia learning. In N. Rummel, M. Kapur, M. Nathan, & S. Puntambekar (Eds.), *To See the World and a Grain of Sand: Learning across Levels of Space, Time, and Scale: CSCL 2013 Conference Proceedings Volume 2 — Short Papers, Panels, Posters, Demos & Community Events* (Vol. 2, pp. 149–152). Madison, WI.

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning*, *4*(1), 33–45. https://doi.org/10.1007/s11409-008-9031-3

Schraw, G., Kuch, F., & Gutierrez, A. P. (2013). 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

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Soderstrom, N. C., Clark, C. T., Halamish, V., & Bjork, E. L. (2015). Judgments of learning as memory modifiers. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 553–558. https://doi.org/10.1037/a0038388

Soller, A., Martínez, A., Jermann, P., & Muehlenbrock, M. (2005). From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. *Int. J. Artif. Intell. Ed.*, *15*(4), 261–290.

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*(1), 204–221. https://doi.org/10.1037/0278-7393.1.204

Stegmann, K., Wecker, C., Weinberger, A., & Fischer, F. (2012). Collaborative argumentation and cognitive elaboration in a computer-supported collaborative learning environment. *Instructional Science*, *40*(2), 297–323. https://doi.org/10.1007/s11251-011-9174-5

Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning*, *2*(4), 421–447. https://doi.org/10.1007/s11412-007-9028-y

Suthers, D. D. (2001). Towards a systematic study of representational guidance for collaborative learning discourse. *Journal of Universal Computer Science*, *7*, 254–277.

Suthers, D. D., & Hundhausen, C. D. (2003). An experimental study of the effects of representational guidance on collaborative learning processes. *Journal of the Learning Sciences*, *12*(2), 183–218. https://doi.org/10.1.1.113.5739

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, *10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

Tenney, E. R., Small, J. E., Kondrad, R. L., Jaswal, V. K., & Spellman, B. A. (2011). Accuracy, confidence, and calibration: How young children and adults assess credibility. *Developmental Psychology*, *47*(4), 1065–1077. https://doi.org/10.1037/a0023273

Thiede, K. W. (1999). The importance of monitoring and self-regulation during multitrial learning. *Psychonomic Bulletin & Review*, *6*(4), 662–667. https://doi.org/10.3758/BF03212976

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, *95*(1), 66–73. https://doi.org/10.1037/0022-0663.95.1.66

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: An analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 1024–1037. https://doi.org/10.1037/0278-7393.25.4.1024

Valcke, M. (2002). Cognitive load: Updating the theory? *Learning and Instruction*, *12*(1), 147–154. https://doi.org/10.1016/S0959-4752(01)00022-6

Vernon, D., & Usher, M. (2003). Dynamics of metacognitive judgements: Pre- and post retrieval mechanisms. *Journal of Experimental Psychology, Learning, Memory and Cognition*, *29*(3), 339–346. https://doi.org/10.1037/0278-7393.29.3.339

Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science*, *33*(1), 1–30. https://doi.org/10.1007/s11251-004-2322-4

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, *46*(1), 71–95. https://doi.org/10.1016/j.compedu.2005.04.003

Wilcox, R. (2012). *Introduction to robust estimation and hypothesis testing* (3rd ed.). Boston: Academic Press. https://doi.org/10.1016/B978-0-12-386983-8.00020-2
Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Lawrence Erlbaum.
Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning*, *12*(4), 423–467. https://doi.org/10.1007/s11412-017-9267-5
Yates, J. F., Price, P. C., Lee, J.-W., & Ramirez, J. (1996). Good probabilistic forecasters: The ‘consumer’s’ perspective. *International Journal of Forecasting*, *12*(1), 41–56. https://doi.org/10.1016/0169-2070(95)00636-2