Social Metacognition in Dyadic Learning

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Michael Schubert, M.Sc.
aus Lippstadt

Tübingen
2016

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	22.05.2017
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Dr. Friedrich W. Hesse
2. Berichterstatter:	Prof. Dr. Daniel Bodemer

Table of Contents

Table of Figures

List of tables

Zur besseren Lesbarkeit wird im Text für alle Personen die weibliche Sprachform gewählt. Selbstverständlich sind, wenn nicht anders erwähnt, beide Geschlechter gemeint. Die Wahl der weiblichen Form beinhaltet keine Wertung.

Chapter 1

Introduction

Cooperative learning, and especially dyadic learning as its most basic form, is, by now, a well-established pedagogical means almost every one of us experienced back in school. Whether we freely grouped together to prepare for an upcoming exam, or whether we were instructed by the teacher to cooperate in project-based or problem-based learning tasks (Aronson, 1997; Barron et al., 1998), cooperative learning was and still is an omnipresent phenomenon.

With the rise of virtual learning platforms and digital classrooms (i.e. flipped classroom concept, Bishop & Verleger, 2013; Stone, 2012; Strayer, 2012), learning will be orchestrated more than ever with a strong focus on group interactivity while the acquisition of new knowledge will be transferred to individual learning phases, e.g. as homework. It is likely that with more interactivity in the classroom, the amount of cooperative learning tasks will increase. Cooperative learning requires learners to show strong regulation skills: groups have to coordinate their plans, strategies and activities in order to solve the problem and to reach mutual learning goals (Winne, Hadwin, & Perry, 2013).

But at the heart of every form of cooperative learning stays the actual knowledge exchange. When preparing for an exam, learners first need to individually acquire new knowledge – new facts, relations and functions among these facts. Then, they can elicit crucial parts of the previously acquired learning material from the other person, discuss not-understood items or clarify misconceptions within the group. In such learning sessions for the preparation of an exam, learners mostly benefit from mutual discussions by explaining their understanding of the learning material to the partner, who in turn can respond and articulate their understanding of the learning material (O'Donnell, 1996; Palincsar, Brown, & Martin, 1987; Webb, 1989, 1991). When teachers implement cooperative learning methods in class, they try to create similar situations. For example, they might distribute complementary text material with regard to a problem-solving task among members of a group, so that each member becomes an expert of her learning material (cf. *jigsaw* method; Aronson, 1997; *positive resource interdependency* in cooperative learning; Johnson, Johnson, & Smith, 2007). Afterwards, learners have to integrate their acquired knowledge of the learning material with those of other group members in order to create a complete understanding of the overall material within the group, which enables them to finally solve a part of the learning task.

Both examples have one crucial aspect in common: through the course of a first learning phase *differences of understanding* arise between group members. Somehow, these differences are recognized by each group member and serve as a trigger forcing the group to balance each other's level of understanding. But how are differences regarding the level of understanding cognitively represented, and what are the cognitive mechanisms which help individuals in such group learning phases to decide whether to provide or to demand for relevant information?

From cognitive psychology it is known that learners are capable of representing their knowledge[1] and (simultaneously) reflect and influence their knowledge states through cognitive processes of monitoring and control, which has been termed *metacognition* (Flavell, 1979; Hacker, 1998). However, studies in this research field concentrate solely on scenarios of *individual* learning, while *group* learning scenarios have been neglected so far. The concept of *group knowledge and information awareness* tried to cover that research gap (Buder & Bodemer, 2008; Dehler, Bodemer, Buder, & Hesse, 2011; Engelmann, Dehler, Bodemer, & Buder, 2009; Engelmann & Hesse, 2010), but did not specify relevant cognitive instances and mechanisms, which could explain individual learning behavior in group learning situations. And answers to how learners might represent their learning partners' knowledge are not trivial. Research on social cognition demonstrated, for example, that the representation of others' cognitive states is biased, but becomes more accurate the more information the subject receives (Fussell & Krauss, 1989; Ladegaard, 1995; Letzring, Wells, & Funder, 2006). Moreover, once differences in the level of understanding are recognized, people benefit from trying to resolve that differences. Literature on small and cooperative group learning have demonstrated that both explainers and 'explainees' benefit from elaborated explanations (Lou et al., 1996). Additionally, both learners benefit when explainers achieve to respond precisely to explainees' requests (Webb, 1989). Furthermore, scholars in the realm of cooperative learning studied the antecedents, which lead learners to engage in

[1] In the remaining part of this dissertation I will use the word 'knowledge' synonymously to a learner's understanding of the learning material, and define understanding as the linkage between facts (e.g. the human body contains a small and a large intestine, compare Study 1) and their relationships or functions among each other (e.g. through contraction and relaxation of the surrounding muscles, nutrition is transported from the small to the large intestine).

helping and encouraging each other (Johnson et al., 2007; Johnson & Johnson, 1999). They showed that different but complementary knowledge (positive interdependence; Buchs, Butera, & Mugny, 2004; Ortiz, Johnson, & Johnson, 1996) is needed to engage in promotive interactions creating a mutual understanding of the task at hand. Currently, researchers try to better understand the actual regulative interactions within the group, the way how learners negotiate mutual goals, plans and strategies, and how they express acts of monitoring and control. (Goos, Galbraith, & Renshaw, 2002; Iiskala, Vauras, Lehtinen, & Salonen, 2011; Khosa & Volet, 2014; for frameworks of micro- and discourse analysis in group regulation see; Hadwin, Järvelä, & Miller, 2011; Winne, Hadwin, & Perry, 2013). Nonetheless, whether knowledge differences are cognitively represented by each group member, whether they are cognitively monitored and whether they are linked to regulative interactions is unclear.

From a communication perspective, peer and promotive interactions can be interpreted as a way of how groups regulate different knowledge states between individual group members. However, while one can argue that learners must reflect on the specific knowledge constellation in order to efficiently interact with their group members (Lou et al., 1996, 2001; Webb, 1991), to my knowledge no model exists that describes the individual cognitive perspective and the way each learner represents, monitors and regulates own *and* other group member's knowledge. Certainly, such a model could help to make cooperative learning more effective. With such a model in mind, cognitive and instructional psychologists, as well as scholars and practitioners from the learning sciences, might better understand and support group learning behavior, e.g. by creating more effective instructional or technological group learning environments. In responding to this need, the current dissertation will develop and partially test a cognitive model of knowledge representations and regulation in cooperative learning situations of dyads. Thus, the cognitive model likes to contribute to discussions of how students in learning groups do and should regulate each other in order to achieve the best possible learning gains.

Addressing this cognitive regulatory view of learners, the model will be based on theoretical frameworks in metacognition (Flavell, 1979; Hacker, 1998; see section 1.1), and on findings in social psychology, which can be re-interpreted under the label of *social metacognition* (Jost, Kruglanski, & Nelson, 1998; see section 1.2). Inspired by their seminal paper in 1994, the model, termed MEDIA (Metacognition in Dyadic InterAction), represents an extension of Nelson and Narens' metacognitive model (see section 1.3) in that it describes how learners in a dyadic learning situation represent, monitor and regulate their own and their partner's knowledge. The core of the model assumes a reciprocal influence between monitoring and control behavior. According to the model, monitoring own and partner understanding of a learning material has an influence on learner's control behavior (e.g. providing and eliciting explanation, mutual elaboration), while, conversely, given or received explanations during interaction have an influence on monitoring own and partner knowledge representations. After describing the model, I present three separate empirical studies demonstrating the range of validity of MEDIA. Testing the influence of monitoring on control behavior, the first study investigated participants' expectations and intentions concerning their peer interaction under various, cognitively represented differences of understanding (see section 2). Confronted with a fictitious learning partner with a similar or different level of understanding after a first individual learning phase, participants were asked to represent and monitor own, as well as partner knowledge states, and to indicate their potential control behavior for a collaborative learning phase. Findings suggest that participants take absolute knowledge differences (differences over all learning items), but also to some extent relative differences (item-by-item differences) between own and partner level of understanding into account when indicating their control behavior. As the results of the first study emphasize the interactional potential of heterogeneous knowledge constellations within a dyad, the second study tested further assumptions of the model in the interplay of low- and high-self-judged learners during a real collaborative learning phase (see section 3). Besides confirming some results from Study 1, Study 2 could further demonstrate that the accuracy of monitoring one's own understanding can be transferred to more accurate regulation behavior within the dyad. Surprisingly, this did not lead to higher learning gains for the partner. Nonetheless, learners' control behavior within the group had some effects on partner-monitoring accuracy, while self-monitoring accuracy was not affected. More specifically, receiving explanations was related to judging the partner, but providing explanations showed no effects on judging one's own knowledge. However, a strong false consensus bias was found, leading learners to presume their partners exhibited the same level of understanding as they had themselves. In order to investigate whether collaborative elaborations also have an influence on monitoring behavior in a problem-solving scenario, Study 3 tested dyads which had either identical or complementary prior knowledge with regard to a problem-solving task (see section 4). Even though the knowledge constellation had no influence on the problem-solving performance, results emphasized that learners are heavily influenced by their own knowledge state when judging a partner's knowledge state—thus confirming the findings of a false consensus from Study 2. The dissertation concludes with a general discussion including limitations to the developed model, implications for different research fields and future research directions (see section 5).

1.1 Metacognition

In order to provide the basis for a cognitive perspective on regulatory processes in learning groups, one has to describe how learners actually represent their own knowledge, and how they are able to mentally reflect about their own cognitive states. While theories exist that describe how knowledge is created and represented in working and long term memory (Baddeley, 1986; Mayer, 2005; Moreno & Mayer, 2007), metacognition focuses on the question if learners are able to simultaneously reflect and influence these mental representations of knowledge. Although the term metacognition has been used to describe slightly different, but nonetheless related concepts in the past, it can be defined as 'thinking about thinking' or 'knowledge and cognition about one's own thoughts' (Flavell, 1979; Hacker, 1998; Nelson & Narens, 1994). These definitions of thinking can be understood as what learners know about their own cognitive system and functioning (i.e. metacognitive knowledge), as the way learners get aware and reflect about their own cognitive or affective state (i.e. metacognitive experiences), and as what learners know about strategies to influence own thinking (i.e. metacognitive skills). Depending on how learners reflect and evaluate their learning process, hence how they monitor their own cognitive state (which become often manifest in metacognitive experiences), metacognitive skills appear as a specific control behavior in the learning context. Therefore, metacognition often refers to metacognitive monitoring and metacognitive controlling of learner's cognitive states (e.g. knowledge, comprehension, attitudes; Hacker, 1998).

Building on a control-process-view (Carver & Scheier, 1990), Nelson and Narens' (1994) seminal paper illustrated the idea of monitoring and controlling one's own thoughts through an object-level and a meta-level. On the object-level, various cues are processed allowing the learner to represent her understanding of the world, thus creating a cognitive state of the world. This cognitive state itself can then be the target of simultaneous cognitions on the meta-level. In the context of learning, information is processed by reading learning material and is then represented on the object-level as the current state of knowledge (i.e. facts and relations among these facts). The meta-level in turn contains both a desired state of knowledge (i.e. the learning goal of the learner) and possible actions (i.e. learning strategies) to potentially change the processing on the object-level. Nelson and Narens (1994) postulated that object-level and meta-level mutually inform each other by way of two flows of information. The flow of information from the object-level to the meta-level brings the current state and the desired state together, thus leading to an evaluation of one's learning progress. This flow of information is referred to as *monitoring*. The second flow of information moves from the meta-level back to the object-level: e.g., if learners have not met their goals yet, they can control their cognitive behavior by changing their memorizing or reading strategies in order to gradually reduce the discrepancy between current state and desired state (i.e. discrepancy-reductionist-model; Dunlosky & Hertzog, 1998). Therefore, this flow of information is referred to as *controlling*.

As monitoring and controlling build a loop that connects object-level and meta-level, both processes are positively interdependent from each other. Therefore, it is generally held that accuracy in monitoring leads to accuracy in controlling, and vice versa. For instance, if learners accurately assess which parts of the learning material they understood well and which parts they did not understand so well, they can better allocate their resources to non-understood material (e.g., Nietfeld, Cao, & Osborne, 2005; Thiede, Griffin, Wiley, & Redford, 2009; Winne & Perry, 2000). Hence, learners who monitor themselves more accurately will be more efficient in regulating their learning and increase their learning achievements compared to learners with poor monitoring (Ariel, Dunlosky, & Bailey, 2009; Metcalfe, 2009; Nietfeld & Schraw, 2002; Soderstrom & McCabe, 2011; Thiede, Anderson, & Therriault, 2003). However, an emerging view in metacognition is that controlling also affects monitoring (Koriat, Ma'ayan, & Nussinson, 2006; Koriat, 2012). For instance, when learners spend more time on learning an item compared to other items, they might use this information as a cue to lower their judgment of comprehension (and recall) for this learning item. Consequently, the more cues a learner possesses with regard to a specific learning item (besides time, e.g., the ease of processing or the coherence of the text), the more accurate the monitoring will be.

The concept of metacognition entails that, on the object-level, learners construct a representation of their own knowledge, their self-representation, while processing learning material. Learners then cognitively regulate their self-representation through processes of monitoring and controlling via a meta-level, which contains a goal state and possible strategies to interfere. According to several researchers, these processes can mutually influence each other.

While the overwhelming majority of metacognition research deals with individual information processing, back in 1982, Kluwe already proposed that not only "the subject has some knowledge about his own thinking", but also about "that of other persons" (Hacker, 1998). Thus, people can think about other's mental states, which can be called other-representations (i.e. social understanding; Astington & Jenkins, 1995; Carpendale & Lewis, 2004). Similarly, from the field of CSCL (computer-supported collaborative learning) it is known that holding an awareness of the knowledge states of other group members is generally possible and useful for subsequent group communication and the coordination of group activities (*group awareness*, *knowledge and information awareness*, Engelmann, Dehler, Bodemer, & Buder, 2009; Engelmann & Hesse, 2010). While using different digital group awareness tools proved that usefulness (Buder & Bodemer, 2008; Dehler et al., 2011; Sangin, Molinari, Nüssli, & Dillenbourg, 2011), the exact individual cognitive mechanisms behind stayed unclear. So the question arises how monitoring and controlling processes concerning other-representations might function in social learning situations.

1.2 Social Metacognition

Transferring assumptions from metacognition to the social realm, some questions about monitoring and controlling can be asked. In accordance with individual metacognition, there is little doubt that individual learners will also monitor their own understanding when they are in a group learning scenario. But will they also build up and monitor a representation about the knowledge state of their learning partner(s)?

According to Jost et al. (1998), many classical social psychological findings could be re-interpreted in terms of social metacognition, and they conclude that humans indeed also form mental representations about their interaction partners. Little surprisingly, these representations about others are rarely accurate, but rather tend to be biased; a phenomenon that is called false consensus bias (Birch & Bloom, 2004; Marks & Miller, 1987). The authors describe that phenomenon as the personal belief of an individual that others, often unknown people, supposed to think or act as themselves, thus searching for consensus in their thoughts about others (a description how the false consensus bias might manifest cognitively, can be found under the term knowledge imputation, Nickerson, 1999). The false consensus bias has been found in a variety of different fields. For example, highly professional athletes generally overestimate the use of recreational drugs in other sport disciplines compared to their own discipline, especially when they show a strong drug use history by themselves (Dunn, Thomas, Swift, & Burns, 2012). A similar transfer from own experiences to the judgment of others has been found in solving word anagrams. Those participants who solved word anagrams in a first phase, judged these anagrams in a second phase as easier to solve for others compared to new anagrams and compared to participants who did not solve these anagrams in a first phase (Kelley & Jacoby, 1996). Without self-experiences, judgments seemed to rely more on theories and rules about the specific word anagram at hand. Therefore, researchers assume that without appropriate information of the target person or the population, people follow an egocentric view basing their judgments on own knowledge and experiences.

However, there is also evidence that other-representations become more accurate the more information people receive about the target person or the target population (Barkhuysen, Krahmer, & Swerts, 2005; Fussell & Krauss, 1989; Ladegaard, 1995; Swerts & Krahmer, 2005). Moreover, when people interact with the target person they use the exchanged information to adjust their judgments about their interaction partner. For example, Letzring, Wells and Funder (2006) found that both quantity and quality of exchanged information in an unacquainted triad influenced accuracy of other-representations with regard to several personality factors. Similar effects have been found for different degrees of acquaintanceship, suggesting that the longer dyads knew each other (i.e. the more information they possessed about the partner) the higher their accuracy of other-representations (Biesanz, West, & Millevoi, 2007; Watson, Hubbard, & Wiese, 2000). It seems that the more familiar people get with their partners the more attracted they are to them, which is mediated by partner's responsiveness and their perceived knowledge of the partner (Reis, Maniaci, Caprariello, Eastwick, & Finkel, 2011). In other words, the more partners interact the more they get to know each other, and the more they become emotionally close to the partner. This in turn can cause a higher willingness to help (Korchmaros & Kenny, 2006).

Applying these research findings to dyadic learning situations, it can be hypothesized that learners might be capable of holding and judging other-representations. In the presence of an actual learning partner, the representation might be called partner-representation. If no other information is available learners might base their partner-judgments on self-representations of knowledge and experiences. But the more cues a learner has available about the learning partner, the more accurate the judgments about the partner might become, which is in correspondence with learner's tendency to base the monitoring of one's own understanding on available cues during learning (see section 2.1; Koriat, 1997). Finally, the more learners get to know each other the more they might help each other. Therefore, the literature suggests that interacting might be seen as a sort of control behavior in dyads, first in order to get to know each other, and second to interact in a way both learners can reach their personal goals – this can simply mean to grasp a partner's understanding, or to complete a mutual learning goal. But how does interaction and information exchange manifest?

Research literature on small group and cooperative learning suggests that mutual explanations might play this part and that they are important triggers of achievement (Cohen, 1994; O'Donnell, 1996; Palincsar et al., 1987; Webb, 1991). From a cognitive elaboration perspective, it is argued that elaborations change cognitive structures of learners (Slavin, 2011). A means to force students to elaborate is to let them explain the learning material to their peers. For instance, O'Donnell and colleagues (1996) proposed *scripted cooperation*, a method to enhance students' learning achievements on text comprehension. After reading one passage of a given text material, one learner of the dyad takes the role of the explainer while the second learner has to listen and detect errors or omissions. For the next passage, learners switch roles. A similar method is the *reciprocal teaching* method (Palincsar et al., 1987). In classroom, student tutors are designated by their teachers, assigned to small groups and instructed with strategies (i.e. questions generating, summarizing and clarifying) to elicit text comprehension in their tutees.

Substantial evidence exists for the effectiveness of both methods (for reviews on scripted cooperation see Dansereau, 1988; O'Donnell & O'Kelly, 1994; for effects of reciprocal teaching see Palincsar & Brown, 1984; Spörer, Brunstein, & Kieschke, 2009). While dyads instructed with scripted cooperation achieved higher comprehension rates than individuals (Dansereau, 1988; O'Donnell & O'Kelly, 1994), a deeper look into the different roles revealed that explainers in dyads seem to benefit more from the method (when scripted cooperation was not implemented with switching roles). This beneficial effect for explainers go hand in hand with Webb's (1991) review of mathematical problem solving in small learning groups. Her results underscore that providing explanations, cues and hints helped explainers to foster their understanding, while she pointed to the fact that learners not always benefit from received information, especially if just the solution to potential test questions were provided. However, other reviews suggest that both explainers and explainees gain their understanding of the learning material comparably well during cooperation (Lou et al., 1996, 2001).

In the field of cooperative learning, researchers tried to find the antecedents leading to fruitful peer interactions of giving and receiving information, or, how Johnson et al. (2007, 1999) formulated it, to *promotive interactions*. The most important antecedent is *positive interdependence*. In contrast to negative interdependence, which promotes competition, positive interdependence is said to encourage members of a group to cooperate. Positive interdependence among group members can be established differently, e.g., through complementary resources or simply through spatial proximity, but the most common way is by creating positive *outcome* interdependence, which is giving the group a shared goal. Only if all members reach their goals, the group can be successful. Generally, positive outcome interdependence leads to more promotive interactions of providing and receiving information, as well as to higher learning achievements (Ortiz et al., 1996), but these achievements depend on the quality of interactions (Buchs et al., 2004). However, even if rules are introduced to promote high-quality discussions (e.g., in cooperative controversy), increased support, question generation and provided explanations do not necessarily lead to higher learning gains (Golub & Buchs, 2014).

As can be seen, in small group and cooperative learning literature interactions among group members are essential for successful learning, but the effects on learning achievement are ambiguous. The heterogeneity of the learning group together with a positive outcome interdependence seem to be a good trigger for learning groups to engage in more question asking and more explanation giving, but providing or receiving explanations not always leads to more learning gains, neither for explainers nor for explainees. And even when the interactions are highly structured like in cooperative controversy, higher learning gains are not guaranteed. Therefore, a better understanding of the cognitive processes is necessary to systematically analyze barriers and resources of the individual learner, as well as of the dyadic setting.

Despite their ambivalent impact on learning achievement, one could assume that providing explanations, eliciting explanations through questions, and receiving explanations are the principle actions of social regulation in small and cooperative group learning (Webb, 1989, 1991). On the one hand, when a learner explains certain parts of the learning material, this will foster her understanding and makes her more self-aware of what she understood well or not so well. Asking questions works in a similar way. Learners have to identify the parts of their understanding for which they need clarifications, and therefore they get more aware of their own understanding. On the other hand, receiving explanations often contains new information or new conclusions and hence enables learners to increase their understanding. In the next section, I will lay out in more detail the cognitive mechanisms of provided and received explanations using the object-meta-level model, while I embed my considerations into a dyadic learning scenario under positive outcome interdependence.

1.3. MEDIA: MEtacognitive model in Dyadic InterAction

Based on Nelson and Narens' (1994) model described above, I present an extension that seeks to explain how learners monitor their own *and* their partner's understanding of a given learning material, while it also gives insight into how providing and receiving information functions as a control behavior in dyads. Moreover, the extended model explains how control behavior influence learning achievement and monitoring of self and partner. The extended model, termed MEDIA (MEtacognitive model in Dyadic InterAction) is displayed in Figure 1.

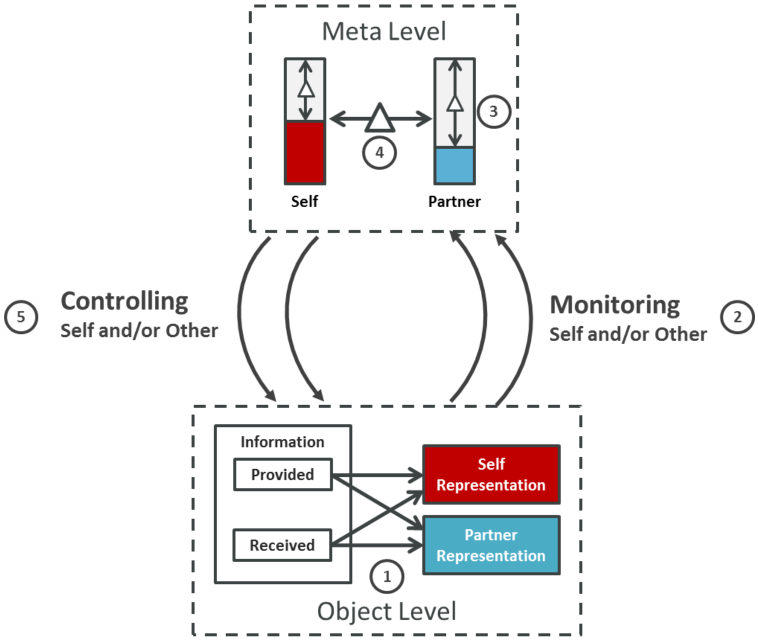

Figure 1. MEtacognitive model in Dyadic InterAction
An Extension of Nelson and Naren's (1994) object-meta-level model to dyads. 1. Learners process provided or received information through dyadic interaction: provided information creates more accurate self-representation of knowledge and can lead to partner learning, while received information creates more accurate partner-representation of knowledge and can lead to better self-learning. 2. Monitoring of learner's own and/or partner's current state of knowledge, 3. Individual discrepancies, 4. Intragroup differences. 5. Controlling of learning according to individual discrepancies and intragroup differences either through explanations or questions

Like Nelson and Narens' (1994) model, MEDIA distinguishes between an object-level and a meta-level. However, on both levels not only a learner's self-representation, but also her partner-representation is taken into account. Starting at the bottom of Figure 1, the creation and updating of both a self-representation and a partner-representation on the object-level is achieved through the dyadic exchange of explanations. First, by receiving explanations from the partner, a learner's own understanding is updated. Moreover, the received information (irrespective of whether it is an explanation or a question) is an important cue that informs about the knowledge state of the partner. Second, by providing explanations to the partner, a partner understanding is updated. Moreover, the provided information (irrespective of whether it is an explanation or a question) serves as a cue informing about the knowledge state of the self. This latter assumption builds on findings from the metacognition literature which show that "generating" one's knowledge before judging one's knowledge is associated with higher monitoring accuracy (*generating effect*; Dunlosky & Lipko, 2007; Thiede et al., 2003). Summing up the assumptions about the object-level, receiving questions or explanations leads to better monitoring accuracy of the partner (with the latter improving one's learning). In contrast, providing questions or explanations leads to better monitoring accuracy of the self (with the latter improving the partner's learning).

Once they are formed, self- and partner-representations can be transferred from object- to meta-level. Here, I suggest that learners hold two goal states in mind: learner's own desired state of knowledge and partner's desired state of knowledge (e.g., task goals). In general, these goal states are independent of each other. But in the context of positive outcome interdependence, hence shared goals, own and partner's desired states become interdependent. In most cases this means that own and partner goals are equalized, e.g. when both learners have to achieve the same percentage score in a performance test, or when they have to combine prior knowledge to solve a problem. However, learners can evaluate *individual discrepancies* between own representation and own goal, and between partner representations and partner goal separately. Moreover, learners can assess *intragroup differences* between own and partner representation of knowledge. If cues from interaction are sparse (e.g., at the beginning of the information exchange), false consensus bias might first lead to a high similarity between the evaluated self-representation and the evaluated partner-representation, but the more cues are drawn from the partner (because learners need to better estimate how much the partner actually knows), the representations become more salient. Of course, learners will try to reduce their own individual discrepancy (Dunlosky & Hertzog, 1998), but in dyadic learning they also need to take care of their partner's individual discrepancy and therefore take the intragroup difference into account. No difference will cause a balanced amount of provided explanations (help-giving) and elicited explanations (i.e. asking questions: help-receiving). But if a positive difference occurs between own and partner understanding, learners are more likely to provide explanations; if the difference is negative, learners are more likely to elicit explanations.

MEDIA suggests a reciprocal influence between monitoring and controlling. Monitoring self- and partner-representations will influence the amount of provided or elicited explanations during cooperation. Moreover, if the accuracy of monitoring is high, learners will be more efficient in selecting those parts of the learning material for which explanations need to be exchanged. Conversely, providing or receiving information will influence learning and the accuracy of monitoring of self- and partner-representations. While providing cues might help learners to monitor their self-representations more accurately and help partners to gain a better understanding, receiving cues might help learners to monitor their partner-representation more accurately and gain a better understanding for themselves.

1.4 Research Questions and Overview of Studies

Small group and cooperative learning provides evidence that promotive peer interactions are a key to successful learning. The literature further suggests that learning plays out through a give and take of explanations during interaction, which can be interpreted as a regulative mechanism to balance knowledge differences within the group. From a cognitive perspective, it is proposed that learners need to hold and evaluate both self- and partner-representations of knowledge to decide whether explanations have to be provided or elicited. Further, I assumed that provided and received information take specific functions which influence learning and learner's self- and partner-representations. This has been condensed in the MEtacognitive model for Dyadic InterAction (MEDIA) which can be assigned to the field of social metacognition.

The model contains multiple assumptions which all deserve a closer analysis and examination. In this dissertation, I was especially interested in the question if learners in different dyadic constellations concerning the level of understanding also show different interaction behavior. Consequently, I take a closer look at help-giving and help-receiving actions of dyads under varied knowledge constellations within the dyad, and I examine whether these actions really have an influence on learner's self-and partner monitoring accuracy. At the same time, other factors like the goal states for learners or group size were kept constant. Three empirical studies have been conducted to examine these questions, which are represented by chapter 2-4. Each chapter contains a theoretical and a methodological section, as well as results and discussions, as they were composed as discrete scientific papers. Even though it was taken care to reduce redundancies between the introductory chapter 1 and consecutive empirical chapters, some overlap with regard to content was unavoidable. Apart from formal and content-wise commonalities, I will give a short overview about the main differences and results between each chapter, and indicate to which journal the corresponded papers have been submitted.

- **Chapter 2 based on:**
  Schubert, M., Buder, J., & Hesse, F. (submitted). Social metacognition and heterogeneity in dyadic learning. *Journal of Educational Psychology*

  According to MEDIA, it is predicted that processes of providing and receiving explanations through social interaction is associated with (a) the discrepancy between current and desired states of self and the partner knowledge; and (b) the intragroup difference between self and partner knowledge. The empirical study investigated (1) the effect of low and high self-judged knowledge in combination with either low or high ostensible partner knowledge on the intention to provide and the expectation to receive explanations; and (2) the effect of learner's discrimination abilities of identifying well and poorly understood material (i.e. relative monitoring accuracy) on the intention to regulate the group efficiently (i.e. relative control accuracy). Results on the first question revealed that participants intend to provide explanations mainly based on the amount of own knowledge, while their expectations to receive explanations depended solely on the amount of partner knowledge. Concerning question two, no dependency was found between monitoring accuracy and group regulation accuracy.

Even though the hypotheses of Study 1 have not been fully verified, the outcomes provide evidence for a high learning potential in heterogeneous learning groups. Thus, if a learner judged herself as highly knowledgeable (high-ability), and is matched with a learner who judged herself as less knowledgeable (low-ability), intentions and expectations of both learners can be fulfilled. Similar potentials of heterogeneous grouping have been suggested elsewhere (Saleh, Lazonder, & de Jong, 2005, 2007; Webb, 1991), as high-knowledgeable learners' explanation abilities can cause both internal elaborations at learner's side and knowledge gains at partner's side. However, the first study provides no evidence for a relationship between monitoring and control accuracy, as it was suggested by MEDIA. Therefore, the second study was dedicated to further reveal effects of monitoring on controlling and vice versa. Presented in Study 2, the study incorporated a controlled experiment in which these cognitive mechanisms are investigated in the realm of heterogeneous knowledge levels analyzing real dyad interaction. Another facet of heterogeneity is proposed by cooperative learning methods, which explicitly control for complementary prior knowledge among a group of learners (Aronson, 1997; Ortiz et al., 1996). When learners with different prior knowledge are grouped together, they can "teach" and guide each other, especially in tasks for which the different aspects of knowledge need to be combined. Chapter 4 takes a closer look on such a learning dyad. Setting up another controlled experiment with real dyad interaction, the study tries to answer the questions whether complementary knowledge causes dyads to engage in more promotive interactions, and whether these interactions have an influence on monitoring accuracy of self and partner knowledge.

- **Chapter 3 is based on:**
  Schubert, M., Buder, J., & Hesse, F. (submitted). Social metacognition and heterogeneity in pairs of learners: A study on monitoring and controlling self and partner knowledge. *Learning and Instruction*

  To test whether self-monitoring accuracy influences efficient dyadic interaction and whether dyadic interaction influences self and partner monitoring accuracy, an experimental study was set up with 43 dyads heterogeneous in their self-judged knowledge levels. First, results indicate that high-knowledgeable learner's monitoring accuracy predicts efficient help-giving, but from which low-knowledgeable partners could not benefit. Contrary to my expectations, low-knowledgeable partners mainly gained a better understanding through the amount of received explanations, and not through the efficiency of received explanations.

Second, all learners showed low self-monitoring accuracy, but high partner-monitoring accuracy after the cooperation. While help-giving actions could not support learners to achieve a more accurate self-representation, receiving help contributed to learner's accurate representation of partner knowledge. Moreover, the false consensus effect was identified as a strong bias for learners judging partner's knowledge.

- **Chapter 4 is based on:**

  Schubert, M., Buder, J., & Hesse, F. (submitted). Positive resource interdependence on promotive interactions, group performance and individual free recall in the context of scientific and cooperative information problem solving. *Journal of Applied Cognitive Psychology*

  The chapter investigates the additional effect of positive resource interdependence in cooperative information problem solving of dyads with positive outcome interdependence. 65 dyads were confronted with an authentic information problem and received either identical or complementary material prior to the task. As positive resource interdependence is claimed to increase promotive interactions within the group, I expected positive effects on learners' representations of own and partner knowledge. Results showed that dyads with complementary knowledge were more accurate in representing own knowledge, while no differences occurred between dyads with complementary and identical knowledge with regard to partner representation accuracy. In both cases, monitoring was not influenced by own or partner's interaction behavior. However, learners were strongly influenced by own knowledge representations when predicting partner knowledge.

In the final chapter 5 of this dissertation, I will summarize the results of each chapter, elicit commonalities and differences, point to strengths and limitations, and outline implications for different research fields.

Chapter 2

Study 1 - Social metacognition and heterogeneity in dyadic learning

Chapter is based on

Schubert, M., Buder, J., & Hesse, F. (submitted). Social metacognition and heterogeneity in dyadic learning. *Journal of Educational Psychology*

2.1 Introduction

Anne and Beth are medical students jointly preparing for their upcoming exams. In order to achieve their goals, they have to self-regulate both their individual and their collaborative learning processes (Hadwin et al., 2011). While Anne and Beth might jointly set goals and create learning plans, their main focus during collaborative learning sessions will probably lie on the reflection of each other's understanding and on the mutual explanation of facts and concepts. But how do they accomplish this regulation? Is Anne simply following her own learning progress (i.e. in monitoring how much she understood from learning materials and from her interaction with Beth), or is Anne aware of Beth's learning progress, too? Is Anne taking the difference between her own and Beth's learning progress into account to determine whether to give help, or receive help? And what determines an effective information exchange among the two students? The goal of this chapter is to give insight into these metacognitive mechanisms (Flavell, 1979; Nelson & Narens, 1994) of individual learners during cooperative learning with different knowledge levels.

Metacognition

In section 1.1, metacognition has been described as 'thinking about thinking' (Flavell, 1979; Hacker, 1998). Further, it was explained that 'thinking about thinking' can mean to evaluate one's current state of knowledge, which is often termed *monitoring*, while it can also mean to change one's way of thinking by applying different thinking strategies, e.g., in learning. The latter kind of thinking is therefore often referred to as *controlling* (Hacker, 1998). Nelson and Narens (1994) structured the processes of monitoring and controlling by introducing an object-level and a meta-level. On the object-level, the learner represents her understanding of a given learning material. On the meta-level the learner holds a desired state of knowledge (i.e. the learning goal of the learner) and possible actions (i.e. learning strategies). Nelson and Narens (1994) then defined the two flows of information between object- and meta-level as monitoring and controlling. Through monitoring one's current state of knowledge, the learner can evaluate her learning progress on the meta-level by comparing it to a goal state. Depending on the outcome, changes are transferred back to the object-level to control further learning processes. For example, if learners recognize that they have not progressed the way they wanted, they can influence their learning behavior by applying other learning strategies. Thus, they try to reduce the discrepancy between monitored knowledge state and desired goal state (i.e. discrepancy-reductionist-model; Dunlosky & Hertzog, 1998).

The overarching goal in monitoring one's own learning is to be as accurate as possible in order to optimize subsequent control behavior (e.g. Nietfeld & Osborne, 2005; Thiede, Griffin, Wiley, & Redford, 2009; Winne & Perry, 2000). If accuracy is high, learners can work efficiently on not-learned material, whereas poor accuracy might lead to an inappropriate selection of material. One measure, which is extensively explored in metacognition, is relative monitoring accuracy (de Bruin, Thiede, Camp, & Redford, 2011; Dunlosky & Lipko, 2007; Thiede et al., 2003). Relative monitoring accuracy describes how good learners are capable of discriminating between well and poorly understood material on the basis of individual learning items. Thus, the more learners are accurate in discriminating parts of their understanding, the better they are able to select material for an upcoming learning session. Monitoring and controlling learning items are measured on the basis of subjective, so-called judgments of comprehension (JOC), which are compared either to test items (i.e. monitoring accuracy) or to the selection of certain learning items (i.e. control accuracy). Relative monitoring accuracy analyzes therefore the fit of ordering between JOCs of multiple learning items and corresponding test items, whereas relative control accuracy is analyzed by the fit of ordering between JOCs and the selection of learning items after a first test. In the context of dyadic learning, I will further specify how the fit between JOCs and selected items can be extended to mutual learning items.

As demonstrated, for learners it is of crucial importance that they monitor their own understanding accurately, in order to significantly improve their subsequent learning steps. Especially in groups, relative monitoring accuracy of individually learned material might be a crucial factor when it comes to the question which facts and concepts to exchange in a mutual learning session. But is metacognition exclusively related to one's own knowledge?

Social Metacognition and Cooperative Learning in Small Groups

Metacognition in groups has not garnered much empirical attention (for exceptions see Molinari, Sangin, Dillenbourg, & Nüssli, 2009; Sangin, Molinari, Nüssli, & Dillenbourg, 2011). However, as described in section 1.2, several studies from the field of social psychology can be re-interpreted from a metacognition angle (Jost et al., 1998). By investigating judgments of (unknown) target people, many researchers showed that participants are not only capable of forming a distinct mental state of the target person, but also evaluate the target person, even though they tend to be biased by their own knowledge (Birch & Bloom, 2004; Krueger & Clement, 1994; Moore & Healy, 2008; Nickerson, 1999; Ross, Greene, & House, 1977). To give an example in the context of group learning, prior to a collaborative learning phase learners might strongly assume that their learning partners hold the same understanding as the learners themselves, which, of course, can sometimes be a wrong assumption (Birch & Bloom, 2004).

However, judgments of the learning partner become better the more information the learner has available. In section 1.2 it was demonstrated that both quantity and quality of received information can increase the accuracy of judging for instance a partner's personality (Biesanz et al., 2007; Watson et al., 2000). Moreover, exchanging information in natural conversations increases knowledge awareness of the partner (i.e., grounding in groups; Baker, Hansen, & Joiner, 1999; Brennan, Galati, & Kuhlen, 2010). Building on the presented literature, researchers tried to use computer-supported visualizations in collaborative learning situations to foster a more effective knowledge exchange (Dehler et al., 2011; Engelmann & Hesse, 2010; Sangin et al., 2011). By enabling learners to compare their JOCs with partner's JOCs, Dehler et al. (2011) showed that learners considered own knowledge more for help-seeking actions and partner knowledge more for help-giving actions, while results from Sangin et al. (2011) suggest that learners are more sensitized to their own and their partner's knowledge causing more elaborated utterances.

The literature on social metacognition suggests that learners will take their partners into account when exchanging information. Nonetheless, control strategies in social information exchange might look differently depending on how groups are composed. To address this issue, research on small group learning provides many insights, particularly those studies that examined learning interactions and learning achievement of high- and low-ability students (i.e. students with more/less knowledge) in either homogeneous or heterogeneous groups. In her review, Webb (1991) concluded that learners in heterogeneous groups take over teacher-student-roles, in which high-ability learners adopted the role of a teacher and low-ability learners the role of a student. In this constellation, high-ability learners tended to give more explanations, while low-ability learners received more help, compared to average-ability learners in the group, or compared to homogeneous groups. In homogeneous groups of high-ability learners, all group members lowered their activities presuming that they already knew all answers, whereas in homogeneous groups of low-ability learners, group members created a lot of interaction, but without substantial learning gains.

Subsequent studies on group composition and performance have found somewhat different results. For instance, it was reported that low-ability students can benefit from learning in heterogeneous groups and outperform low-ability students in homogeneous groups (Lou et al., 1996). Furthermore, some studies found that high-ability students in homogeneous groups showed similar learning gains compared to high-ability students in heterogeneous groups (Leonard, 2001; Saleh et al., 2005). Additionally, Saleh et al. (2005) and Saleh, Lazonder, and de Jong (2007) reported increased proportions of elaborated acts during social interaction within heterogeneous groups compared to homogeneous groups, and that these elaborations can be evoked by structured collaborations (i.e. introducing special roles and rules) compared to groups under no constraints.

Though these findings provide a slightly ambiguous picture, some conclusions can be drawn. First of all, dyadic learning often plays out in the form of (mutual) explanations in which one learner provides information to her partner, or conversely, a learner receives information from her partner. Second, whether a learner provides or receives information is partly dependent on the ability, with high-ability learners showing a general tendency to provide explanations and low-ability learners showing a tendency to receive information. Third, the “natural” role assignment of explainer and explainee works best when dyadic learners have different levels of ability, i.e. when dyads are heterogeneous rather than homogeneous. However, it seems that also in homogeneous groups learners try to gain more knowledge and show substantial interactions, although they might realize quickly that mutual learning is hampered and inefficient.

To provide insight into cognitive processes of small cooperative learning groups, a framework was proposed that investigates social metacognition, and particularly the effect of different levels of understanding in dyads. For this, I developed an extension of Nelson and Narens' 1994 object-meta-level model to dyads: MEDIA (section 1.3).

MEDIA: Metacognitive model in dyadic interaction

In Figure 1 (page 21) I presented an extension of the object-meta-level model (Nelson & Narens, 1994) to dyads by explaining first how learners monitor themselves *and* their partner. Second, I demonstrated how information exchange functions as a control behavior in groups, and finally how provided and received information can have an impact on self- and partner-representations of knowledge. One of the interesting assumptions of the model is that control actions, especially providing and eliciting explanations, should appear according to evaluated *individual discrepancies* and *intragroup differences*. More specifically, I proposed that learners a) reduce a partner's knowledge discrepancies by providing new information (through given explanations); that b) they reduce their own knowledge discrepancies by receiving information (through eliciting explanations); and c) the relation between providing and receiving information depends on the extent of intragroup differences between perceived self- and partner-representations of knowledge. How learners might exchange information with regard to different intergroup knowledge constellations is one of the research questions in this chapter.

Another interesting assumption of the model was that the more accurate the monitoring of the self and the learning partner is, the more effective the learning interaction of the dyad will be. That is, evaluating own and partner individual discrepancies accurately will lead to accurate assessments of intragroup differences and a more efficient subsequent control behavior. In this respect, control behavior becomes accurate, hence efficient, when learners fit the amount of provided and received information according to the evaluated intragroup difference. Thus, a second research question in this chapter revolves around the relationship between monitoring accuracy and control accuracy, and the way this relationship is influenced by the different knowledge constellations within a dyad.

The Present Study

The present study comprises the first step of a research program to test some of the proposed assumptions of MEDIA. It is conceptually fueled by inconsistencies found in cooperative learning research of small groups concerning the behavioral tendencies to provide or receive information in heterogeneous and homogeneous groups. In order to extend the literature from a metacognitive angle and to find out how learners might provide or expect information in particular dyad constellations, the study uses a paradigm with high experimental control. That is, rather than having actual dyads exchange information (a focus of the two subsequent studies), I put individual participants in a fictitious cooperative learning scenario (see Johnson & Johnson, 1999; Slavin, 1983), provided information about the (fictitious) learning partner (similar to providing partner JOCs; Dehler et al., 2011; Sangin et al., 2011), and merely asked participants about their intentions to provide and their expectations to receive information. This allowed me to test whether a) these intentions and expectations depend on own knowledge, on partner knowledge, or on the intragroup difference between own and partner knowledge; and b) whether accuracy in monitoring one's own knowledge is associated with the accuracy of control intentions. As information about the (fictitious) partner knowledge was provided to the participants, the experimental setup of this study could not address the question of whether there is a relation between accuracy in monitoring partner knowledge and accuracy of control behavior.

Based on the majority of cooperative learning literature presented above, and on the resulting MEDIA model, I expect the following behavioral patterns: When learners with the same learning goal are grouped together, individuals with a high current state of own knowledge (i.e. low individual discrepancy) will have the intention to provide more information than learners with a low current state of own knowledge (i.e. high individual discrepancy). This tendency should be particularly strong in heterogeneous dyads (i.e. when interacting with a low-knowledge partner), because a steeper perceived knowledge slope to partner's knowledge should attract learners to compensate stronger for this inequality than a shallower slope. This results in the following hypothesis:

H1: Learners with a high current state of knowledge have a stronger intention to provide information than learners with a low current state of knowledge, which is especially emphasized in heterogeneous dyads.

Complementary to H1, I assume that learners with a low current state of knowledge will have the expectation to receive more information than learners with a high current state of knowledge, but also only when they perceive to be in a heterogeneous group (i.e. when interacting with a high-knowledge partner).

H2: Learners with a low current state of knowledge have a stronger expectation to receive information than learners with a high current state of knowledge, which is especially emphasized in heterogeneous dyads.

As the literature suggests, high-ability learners might focus more on help-giving actions and low-ability learners more on help-receiving actions. That is, the greater the intragroup difference between learner and learning partner appears, the bigger the focus on the appropriate actions. Therefore, it is very important for learners to discriminate between learning items which they understood very well and those they understood not so well, to adapt their actions accordingly (i.e. relative monitoring accuracy influences relative control accuracy). As high-ability learners focus on the items they know very well, they will be better in adapting their information-providing actions, while low-ability learners will concentrate more on their information-receiving actions in regard to items they do not know so well. Moreover, the bigger the overall intragroup difference, the easier learners might identify the learning items for which they can either provide or receive information, as the knowledge difference of each learning item becomes more salient within the dyad compared to constellations of low intragroup differences.

H3: Learners high in relative monitoring accuracy will show a better ability to accurately indicate for which learning item they want to provide knowledge (i.e. relative control accuracy to provide information) than learners with low relative monitoring accuracy. However, this will only occur if they hold a high current state of knowledge in a heterogeneous dyad.

H4: Learners high in relative monitoring accuracy will show a better ability to accurately indicate for which learning items they want to receive knowledge (i.e. relative control accuracy to receive information) than learners with low relative monitoring accuracy. However, this will only occur if they hold a low current state of knowledge in a heterogeneous dyad.

2.2 Method

Participants and Design

I recruited eighty-eight students (71 female) from a university in Germany for course credit or for a reimbursement of 6 €. Participants were between 18 and 29 of age (M = 22.70, SD = 2.82). The students were coming from a wide range of academic disciplines with a majority of 10 students from psychology. Due to the inability to calculate monitoring accuracy from participant responses, one person's data was omitted from analyses.

Participants were randomly assigned to a 2 x 2 between participants design with current state of participant's knowledge as one factor (low vs. high), and current state of (fictitious) partner's knowledge as the second factor (low vs. high). Participant's knowledge was manipulated by applying a dual-task-methodology (Baddeley, 1986; Brünken, Steinbacher, Plass, & Leutner, 2002) in an individual learning phase. This manipulation was implemented in order to ensure a broad range of individual judgments of comprehension (JOC). Participant's *own average JOC* (self-JOC) was taken as a continuous independent variable for further analysis.

Material

Six interconnected text paragraphs about the "Digestive System of Humans" were extracted from a German biology textbook (Bayrhuber, Kull, & Linder, 2005). Each paragraph contained 172 to 176 words, was exactly ten sentences long, and was given a short and distinct title. For each paragraph, I created two multiple-choice test items with five response options each. In a preliminary study with 24 students from the same cohort I found that the paragraphs were rated as fairly well with regard to legibility (1 = *not legible at all*; 6 = *highly legible*; M = 3.96, SD = 0.70) and comprehensibility (1 = *not comprehensible at all*; 6 = *highly comprehensible*; M = 4.07, SD = 0.62). Moreover, the preliminary study indicated that the average required reading time for each paragraph was about 90s (*M* = 88.5, *SD* = 12.1). Hence, for the main study I used a fixed reading time of 90s per paragraph.

As mentioned before, half of the participants additionally underwent a dual-task procedure. For the dual-task, I looked for a radio report from a national, public radio station (Deutschlandfunk Wissen), which was not related to the text material but nonetheless had some relevance for participants, while fitting the defined time frame. I selected a 9-minute report about how secret service agencies around the world are able to spy and hack into normal internet users' computers, which was cut into six pieces of 90s to match the allocated study time of the written learning material of the parallel task.

Measures

Self-JOC. In order to capture judgments of comprehension, participants had to indicate how well they could correctly answer a comprehension question with regard to each of the six paragraphs of the learning material (on a scale ranging from 0 = *very bad* to 100 = *very well*). The average of these six JOC items was then taken as a continuous variable (self-JOC) for further analyses.

Partner-JOC. In addition to self-JOC, participants received information about the fictitious self-assessments of their learning partners (partner-JOC) for each paragraph. Depending on experimental condition, the partner-JOC was either indicating a learning partner with low knowledge or high knowledge. In the low-partner-JOC (LP-JOC) conditions, the purported partner-JOC for the six paragraphs varied between 15 and 35 points on a scale ranging from 0 to 100 (for an average of 25 points). Conversely, in the high-partner-JOC (HP-JOC) conditions, the partner-JOC for the six paragraphs varied between 65 and 85 points (for an average of 75 points). Partner-JOC was used as a factor for further analyses, coded -0.5 for LP-JOC and +0.5 for HP-JOC.

Monitoring accuracy. Monitoring accuracy for each participant was calculated using gamma correlations between a participant's JOC and his/her individual performance in the test criterion (e.g. Rawson, Dunlosky, & Thiede, 2000; Thiede et al., 2003; for a discussion about the use of gamma in metacomprehension see Gonzalez & Nelson, 1996). Gamma correlations range between -1 and +1, where -1 indicates very low accuracy (i.e. systematically judging poorly understood learning items very high and well understood learning items very low). In contrast, a score of +1 indicates very high accuracy (i.e. systematically judging poorly understood learning items very low and well understood learning items very high). As two test items corresponded to one paragraph judgement, I averaged the test item results of the paragraph to pair them with the corresponding judgment, so that gamma was finally calculated using six pairs of judgments and test scores.

Intentions to provide information / expectations to receive information. Intentions and expectations were surveyed using six questionnaire items, which were adapted for each paragraph of the learning material. An example item for giving information was "How much information about the text will you provide to your learning partner?" (item 3) and an example item for receiving information was "How much information about the text will you receive from your learning partner?" (item 1). For all items, participants used a slider ranging from 0 to 100, with 0 being "nothing" and 100 being "very much". For each paragraph, I calculated Cronbach's alpha for the three providing-information-items and for the three receiving-information-items. Alpha values ranged between 0.709 (for receiving information in paragraph 6) and 0.900 (for providing information in paragraph 6) across all paragraphs. Therefore, separate mean variables for receiving and providing information were computed for each paragraph and overall.

Relative control accuracy for giving / receiving information. To calculate relative control accuracy of learners' intentions and expectations to give or receive information, I took a similar approach to Thiede et al. (2003), but instead of taking solely participant's JOC and matching it to her intentions or expectations, I calculated difference scores between participant's own JOC and partner's displayed JOC for each paragraph. If participants judged themselves higher than their partner, difference scores were positive, indicating that participants could principally provide information on the corresponding text paragraph. In contrast, difference scores turned negative when participants judged themselves lower than their fictitious partner, indicating that they could potentially receive information from the partner. Control accuracy was then computed by calculating Pearson product-moment correlations between the six difference scores and the corresponding mean values from the participants' stated intentions to provide and expectations to receive information. Unlike gamma, Pearson's *r* shows linear correlations within a range of -1 and +1. As one part of this correlation was based on a difference, high control accuracy for giving information would be indicated by a high positive value, whereas high control accuracy for receiving information would be indicated by a high negative value.

Additional variables. I also measured three personality factors, namely *group preferences* (from a self-composed four-item scale), a *Big-5-openness* six-item subscale (Körner et al., 2008), and *social comparison orientation* using a German translation (Jonas & Huguet, 2008) of the original 11-item-scale by Gibbons and Buunk (1999). None of the traits had a significant impact on the analysis, so I excluded them from further report.

Procedure

For each experimental session, up to four participants were invited to the lab, randomly assigned to one of the four conditions, and seated in individual cubicles at a table, where a laptop with headphones was installed. After being verbally briefed, the whole experiment ran on the computer, starting with a questionnaire asking for participant's demographic data and traits (i.e. social comparison orientation, openness from the big-5-scale and group preferences). Following the pre-experimental questionnaire, participants were asked to put themselves into the position of a learner who collaboratively prepares for an exam. I made clear that the learning scenario had two phases: an individual learning phase in which they were given time to study six paragraphs about the human digestive system, and a collaborative learning phase, in which they, in theory, could freely exchange information about the material, but without the possibility to look into the learning material again. It was also made clear that the second phase would not involve an actual information exchange with a real person, but was only meant to capture intentions and expectations. Moreover, in order to create positive outcome interdependence (Johnson et al., 2007), participants were informed that they should imagine having the joint goal to achieve an average of 90 % together with their fictitious partner in a knowledge test. Then, the individual learning phase ensued. Having their headphones on, participants were presented with the first written paragraph on the human digestive system for a duration of 90s (indicated by a little timer in the upper right corner of the screen). All participants were allowed to leave the screen and finish their reading before the 90s ended. If they did not abort earlier, the screen switched automatically once the 90s were reached. Each text screen was followed by the JOC question "How well will you be able to answer comprehension questions concerning the paragraph, when I will ask you in 20 minutes?" (captured by a slider ranging in increments of 1 from 0 = *very bad* to 100 = *very well*). The same procedure of reading a paragraph and answering the JOC question was repeated for the remaining five text passages. During the individual learning phase, participants in the low-self-JOC (LS-JOC) conditions received an additional radio report via headphones, and were instructed to actively listen to the radio report in addition to the text reading task (applying a dual-task-methodology; see Baddeley, 1986; Brünken et al., 2002). Moreover, I announced that a test after the individual learning phase would encompass both questions about text material and questions about audio material. I set the sound volume in a way that it did not exceed 55dB. Participants in the high-self-JOC (HS-JOC) conditions were also instructed to wear headphones, but had no additional task and heard nothing during the first learning phase. After the individual learning phase, a distracter text appeared for 30s to clear working memory. After that, actual test performance was captured. Participants answered to twelve test items (two test items for each paragraph), which were shuffled for each participant into a random order beforehand and presented each on one screen. Contrary to the instructions of the LS-JOC conditions, participants did not need to answer questions concerning the radio report. After the knowledge test, information about self-JOC and (fictitious) partner-JOC per paragraph was presented on the screen (adjacent display). Partner-JOC was varied according to experimental conditions (high vs. low partner-JOC). Participants were instructed to read and memorize own and partner's judgment values for each text paragraph as good as possible. After clicking 'next', information about self-JOC and partner-JOC was removed, and participant intentions to provide information and expectations to receive information were surveyed for each paragraph. Finally, all participants were asked to indicate their cognitive load and their mental effort during reading the text material. Additionally, participants in the low-self-JOC conditions were asked to indicate how much they were distracted by the radio report. Once completed, participants were debriefed and remunerated. Generally, the learning procedure was aligned with Thiede's approaches (de Bruin et al., 2011; Thiede et al., 2003) and was designed in line with a research program including Study 2, in which a comprehension test is incorporated before and after the collaborative phase to capture knowledge gains. The entire duration of the experiment was approximately 45 minutes.

2.3 Results

In order to check whether the manipulation of self-JOC worked, participants in the low-self-JOC conditions (LS-JOC) were asked how much they were distracted by the radio report (by the voice of the host of the radio report: $M = 6.19$, $SD = 1.05$; by the voice of the interviewees: $M = 5.79$, $SD = 1.19$; by other noises: $M = 5.00$, $SD = 1.66$; average: $M = 5.66$, $SD = 0.98$; all on a 7-points Likert scale ranging from 1 = *not distracted at all* to 7 = *very distracted*). Further evidence that participants in the LS-JOC conditions were attending to the radio report was coming from their interest ratings: $M = 5.00$, $SD = 1.76$ on a 7-points Likert scale (1 = *not interesting at all*; 7 = *very interesting*).

In other complementary analyses, it was found that ratings for subjective cognitive load to understand the text material negatively correlated with self-JOC while ratings for perceived mental effort showed no relationship with self-JOC; for cognitive load: $B = -0.02$, $SE = 0.01$, $\beta = -0.31$, $t(86) = -3.006$, $p = .003$, $r = .308$; 1 = *not difficult at all*, 7 = *very difficult*; for mental effort: $B = 0.01$, $SE = 0.01$, $\beta = 0.12$, $t(86) = 1.105$, $p = .272$, $r = .118$; 1 = *not exhausting at all;* 7 = *very exhausting*).

To test the hypotheses, I calculated multiple linear regressions with continuous variables, factor variables and interactions as predictors. Self-JOC and monitoring accuracy have been centered before the analysis, while partner-JOC was dummy coded using -0.5 and +0.5 for low and high levels of partner knowledge. In a stepwise procedure, I controlled all analysis by participants' test results, in case they used their test experiences to calibrate their expectations and intentions of information exchange. As the control variable showed no impact on the regression models, only the first step is reported. Finally, due to partly skewed distributions, significant results were tested using a bootstrapped 95 % CI.

According to Hypothesis 1, participants with high self-JOC should indicate to provide more knowledge than participants with low self-JOC, and this effect should be more pronounced in heterogeneous dyads (i.e., in the condition with low partner-JOC). Therefore, a model was tested to investigate whether the intention to provide information depends on self-JOC, on partner- JOC, or on the interaction between self-JOC and partner-JOC. All three predictors were simultaneously entered into the regression model. Results are shown in Table 1. Even though the interaction term first showed a significant effect, CI revealed only a significant positive main effect of participants' self-JOC on their intention to provide information.

Table 1. Multiple regression model for the intention to provide information. Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and the effect size r. Bootstrapping results are noted in parentheses.

	Δ R^2	B	SE	*β*	p	r
Intercept		51.53	1.19		<.001 [49.29, 53.94]	
Self-JOC		0.54*	0.06	0.66	<.001 [0.367, 0.682]	.678
Partner-JOC		1.89	2.37	0.06	.427 [-2.487, 6.589]	.087
Self-JOC x Partner-JOC		0.27	0.13	0.16	.040 [-0.024, 0.618]	.223
Model	0.50					

Taken together, results for the intention to provide information indicate that participants with high self-JOC have a stronger intention to provide information than participants with low self-JOC. However, against this assumption no pattern of moderation by partner-JOC occurred for the intention to provide information - neither for partners with high JOC (homogeneous dyads) nor for partners with low JOC (heterogeneous dyads). Thus, Hypothesis 1 can only be partly confirmed.

Hypothesis 2 stated that participants with low self-JOC have a higher expectation to receive information than participants with high self-JOC, which will be pronounced in heterogeneous groups, hence, when they are grouped with a fictitious partner with high JOC. Therefore, a model tested the relationship between participants' self-JOC, partner-JOC, and the interaction of self-JOC and partner-JOC on the expectation to receive information. Once again, I simultaneously entered all three predictors into a regression model. Results are presented in Table 2. It can be seen that participants' self- JOC has no relationship to their expectations to receive information, while partner-JOC shows a strong association with the expectation to receive information. Nonetheless, partner- JOC has no moderating effect on the association between self-JOC and expectations.

Table 2. Multiple regression model for the expectation to receive information. Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and the effect size r. Bootstrapping results are noted in parentheses.

	Δ R^2	B	SE	*β*	p	r
Intercept		58.60	1.04		<.001 [56.74, 60.53]	
Self-JOC		-0.06	0.06	-0.08	.289 [-0.211, 0.072]	.116
Partner-JOC		20.04*	2.09	0.73	<.001 [16.26, 24.30]	.725
Self-JOC x Partner-JOC		-0.06	0.11	-0.04	.574 [-0.361, 0.193]	.062
Model	0.51					

Apparently, participants seem to have the expectation to receive information when their partners hold a high level of knowledge and this expectation is irrespective of participants' own average knowledge level. Thus, Hypothesis 2 could not be confirmed.

The remaining analyses use monitoring accuracy as a predictor variable. Monitoring accuracy was expressed as the gamma correlation between the six self-JOC and their corresponding averaged performance values. Looking first at the relationship between self-JOC and monitoring accuracy, the data revealed that monitoring accuracy was negatively related to self-JOC (B = -0.01, SE = 0.00, β = -0.22, t(85) = -2.061, p = .042, r = .218).

Hypothesis 3 held that control accuracy for providing information is best when participants simultaneously exhibit a high monitoring accuracy, have high self-JOC, and collaborate with a partner expressing low JOC (heterogeneous dyad). Thus, I regressed *control accuracy to provide* information on the three predictor variables, as well as on each of the four interaction terms resulting from the three predictor terms (i.e. three two-way interactions and one three-way-interaction). The results in Table 3 suggest that neither any predictor variables, nor the interaction terms had an effect on control accuracy to give information. Thus, Hypothesis 3 could not be confirmed.

Table 3. Multiple regression model for the relative control accuracy to provide information. Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and the effect size r. Bootstrapping results are noted in parentheses

	Δ R^2	B	SE	*β*	p	r
Intercept		0.37	0.05		<.001 [0.253, 0.468]	
Rel. Mon. Accuracy (Rel.Mon.Acc.)		0.15	0.12	0.14	.236 [-0.103, 0.401]	.133
Self-JOC		0.00	0.00	0.09	.447 [-0.005, 0.010]	.086
Partner-JOC		0.01	0.11	0.01	.955 [-0.214, 0.222]	.006
Rel.Mon.Acc. x Self-JOC		0.01	0.01	0.09	.460 [-0.009, 0.020]	.083
Rel.Mon.Acc. x Partner-JOC		0.07	0.25	0.03	.784 [-0.470, 0.496]	.030
Self-JOC x Partner-JOC		0.00	0.01	0.08	.508 [-0.009, 0.019]	.075
Rel.Mon.Acc. x Self-JOC x Partner-JOC		-0.00	0.01	-0.02	.898 [-0.028, 0.032]	.015
Model	-0.05					

Next, in order to test Hypothesis 4, I repeated the same procedure for *control accuracy to receive* information as dependent variable. Therefore, in this regression model I examined the association between monitoring accuracy and control accuracy to receive information, and whether this relationship is moderated by self-JOC, by partner-JOC, or by a combination of both moderator variables. I simultaneously entered all three predictor variables, and the corresponding four interaction terms. In Table 4 it can be seen that none of the predictor terms became significant on the outcome variable. Therefore, Hypothesis 4 could not be confirmed.

Table 4. Multiple regression model for the relative control accuracy to receive information. Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and the effect size r. Bootstrapping results are noted in parentheses

	Δ R^2	B	SE	*β*	p	r
Intercept		-0.17	0.06		.009 [-0.281, -0.042]	
Rel. Mon. Accuracy (Rel.Mon.Acc.)		0.06	0.14	0.05	.667 [-0.198, 0.316]	.049
Self-JOC		-0.00	0.00	-0.01	.960 [-0.007, 0.007]	.006
Partner-JOC		-0.10	0.12	-0.10	.419 [-0.333, 0.137]	.091
Rel.Mon.Acc. x Self-JOC		-0.01	0.01	-0.18	.118 [-0.028, 0.003]	.175
Rel.Mon.Acc. x Partner-JOC		0.19	0.28	0.08	.498 [-0.299, 0.753]	.076
Self-JOC x Partner-JOC		-0.01	0.01	-0.17	.165 [-0.025, 0.004]	.156
Rel.Mon.Acc. x Self-JOC x Partner-JOC		0.00	0.02	0.01	.965 [-0.028, 0.032]	.005
Model	-0.01					

2.4 Discussion

The goal of this study was to test for two of MEDIA's assumptions, namely whether evaluated intragroup differences (caused by monitoring self- and partner-representations of knowledge) have an influence on providing and receiving information, and whether monitoring own knowledge accurately has an influence on controlling the information exchange within in the dyad efficiently (i.e. control accuracy). A tightly controlled laboratory experiment was carried out that put individual learners into a fictitious dyadic learning scenario. As the literature on cooperative learning in small groups shows some inconsistencies about the interaction in different group compositions, I captured intentions to provide and expectations to receive information with regard to fixed states of self and partner knowledge. The results of Hypothesis 1 and 2 point out that participants held both self- and partner-representations of knowledge in mind when they made decisions to provide or receive information to/from their fictitious partners. The higher learners assessed their own learning, the stronger they exhibited an intention to provide information. On the other hand, the higher learners perceived partner's knowledge, the stronger they expected to receive information from their partner. However, contrary to expectations these tendencies were not enforced in heterogeneous dyads. One reason for this potential mismatch could lie in the instructions. Participants were given the overall goal for the dyad to reach an average of 90 % in a learning task, so they might have been equally motivated to provide information to low and high knowledgeable partners—on the one hand to quickly close the knowledge gap of less knowledgeable partners, on the other hand to increase the chance to quickly reach the goal with a highly knowledgeable partner. Moreover, learners have been always motivated to receive information when the partner was highly knowledgeable. Apart from the "student-role" explanation in heterogeneous dyads for low knowledgeable learners (Saleh et al., 2005, 2007; Webb, 1991), an explanation for the behavior of highly knowledgeable learners in homogeneous dyads might be that they see themselves as experts who need to share their knowledge with other experts. Thus, they expect to receive and provide an equal amount of information. Such a pattern would fit the description of expert panels which are implemented in educational scenarios like the jigsaw classroom (Aronson, 1997).

The results provide an interesting perspective to disentangle the inconsistencies found so far in small group learning. Learners seemed to use their evaluations of individual self and partner discrepancies intensively to guide their interaction behavior. As long as the overall learning goal is not yet achieved yet (i.e., speaking in percentiles, in this study this means that on average the individual discrepancies are still greater than 10), learners expect to receive information the more the partner knows, and they have the intention to give information the more knowledge they possess. On the one hand, this confirms reported positive effects of heterogeneous groups (Saleh et al., 2005, 2007; Webb, 1991), but on the other hand also explains positive effects found in homogeneous high-ability groups (Lou et al., 1996; Saleh et al., 2005). Findings in this study add the crucial fact that interactions and learning gains might depend on individual discrepancies, on the (interdependent) goals of the learning group, and whether they have been reached yet or not.

Equally interesting from a social metacognitive perspective, results of Hypothesis 3 and 4 revealed no relationship between participants' relative monitoring accuracy and their relative control accuracy for either providing or receiving information. First, this could have been a matter of the rather small variance in participants' monitoring accuracy. Against the assumptions, the dual-task treatment did not affect relative monitoring accuracy as strongly as it has affected self-JOC. Thus, people were on average quite bad in monitoring their own understanding (M = 0.12), compared to the usual standard found in metacomprehension literature (0.27; Maki, 1998). Instead of using immediate JOC, a delayed JOC after the individual learning might have revealed a greater variance (Rhodes & Tauber, 2011). Second, the statistical association between monitoring accuracy and control accuracy was generally quite low ($\beta_{\text{providing}} = 0.12$; $\beta_{\text{receiving}} = 0.08$): participants achieved high and low control accuracy independently of their monitoring accuracy. One general problem might have been that partner-JOC for each text paragraph were visible only once, namely before participants started indicating their intentions and expectations for each paragraph. Thus, mentally calculating differences between own and partner judgments conceivably was very difficult when JOCs for each paragraph have not sufficiently been memorized. An adaptation to partner knowledge was therefore hampered. Finally, participants might not have perceived the benefits of adapting their control behavior on an item-to-item basis. As the goal was to simply reach 90 % together with the partner in the final exam, there was no need to interpret the goal in a way to reach it also as quickly, or as efficiently, as possible. Rather, they might have assessed and acted just in terms of own and partner's *average* JOC scores.

As to the impact of the results on the assumed model of metacognition in dyadic interaction, there is supportive evidence that learners are able to discriminate and hold both their own current state of knowledge and their partner's current state of knowledge. Further, learners were able to monitor their own and their partner's individual discrepancies, as participants reflected these discrepancies in their intentions to provide or receive information. On the other hand, no clear evidence was found for an influence of intragroup differences on intentions and expectations to provide or receive information. Similarly, no link was found between monitoring accuracy and control accuracy in each of the four conditions. However, Table 3 and Table 4 show that the intercepts for relative group control accuracy are significantly different from zero, which implies that learners have calculated intragroup differences at least to some extent. Nonetheless, the missing correlation between monitoring and controlling is quite surprising, as the literature on metacognition suggests a clear connection between the accuracy of the two processes (e.g., Thiede et al., 2003). Reasons might have been that the very controlled experimental setting and the timing of partner-related information contributed to participant's inability to either adequately monitor themselves or to control the learning group on an item-to-item basis.

Still, I want to emphasize that asking individuals about intentions and expectations in this controlled environment has certain advantages. First, researchers might be interested in how intentions and expectations of information exchange might be altered when the social context changes. A change of the partner description (e.g. changing the belongingness or sympathy), a change of the difficulty to reach the goal (i.e. low or high shared goals), as well as a change of the interdependency of outcome or material (Johnson et al., 2007) might heavily influence participants' responses. These factors might also elicit larger intragroup differences, for example if highly knowledgeable participants have to decide whether they want to share knowledge with high- or low-JOC partners under time pressure (e.g. Kelly & Karau, 1999). Furthermore, studies using fictitious dyads could provide enriched video material about learning partners (Barkhuysen et al., 2005; Swerts & Krahmer, 2005), and learners could be asked to judge their partners by themselves. This way, conclusions could be drawn from the influence of certain cues (e.g. mimic, gestures, and gaze) on these judgements, potentially giving rise to patterns of cue-utilization (see cue-based framework: Koriat, 1997). But also individual differences can influence either judgements about the partner, or the willingness to exchange information, as it has been shown with people under strong social comparison orientation (Ray, Neugebauer, Sassenberg, Buder, & Hesse, 2012). Using facilitation methods (e.g. de Bruin et al., 2011; Rawson et al., 2000), researchers can also try to manipulate participants' monitoring accuracy of own understanding, and potentially the monitoring accuracy of partner understanding, to further study the effect on control accuracy.

Of course, the nature of continuously processing information from the partner, continuously updating current states of own and partner's knowledge, as well as the change of monitoring and controlling both knowledge states, inherently needs real knowledge exchange between two people. Consequently, the next studies in this dissertation will investigate social metacognition among real dyads. If evidence for social metacognition can be found among real learning dyads, future studies might investigate additional monitoring factors. Promising candidate variables might be skill training in group processes (e.g. Nam & Zellner, 2011), the influence of perspective taking skills (e.g. Gockel & Brauner, 2013), or the influence of subtle knowledge awareness tools (Dehler et al., 2011; Engelmann & Hesse, 2010; Sangin et al., 2011). Moreover, researchers might be interested in exhaustive conversation analyses (e.g. Baker, Andriessen, Lund, van Amelsvoort, & Quignard, 2007) to further differentiate how certain utterances during dyadic interaction correspond to own and partner judgments. Having said that, further methodological considerations have to be undertaken concerning the procedure and analysis of real dyadic data (Cook & Kenny, 2005; Kenny, Kashy, & Cook, 2006; West & Kenny, 2011).

2.5 Interlude

Results of chapter 2 suggest that learners might provide explanations based on their evaluated individual discrepancy of own knowledge, while they might elicit explanations based on their evaluated individual discrepancy of partner knowledge. Even though a high proportion of dyadic interaction might occur in both homogeneous high-ability groups and groups of heterogeneous ability levels, the potential for knowledge gains seems to be the highest in heterogeneous groups when low-ability learners can benefit from high-ability learners' elaborated explanations. In contrast to individual discrepancies, the evaluation of intragroup differences could not be fully demonstrated by the first study. Intragroup differences of each learning item became not more salient in heterogeneous groups compared to homogeneous groups, so that learners were better able to discriminate for which learning item they should provide or receive more/less information. However, learners did take intragroup differences of each learning item into account when they indicated their expectations of information exchange for these learning items, even though control accuracy was relatively low (intercepts of: relative control $accuracy_{providing}$ = 0.38, relative control $accuracy_{receiving}$ = -0.18). One of the greatest surprises of chapter 2 was that monitoring accuracy had no impact on control accuracy. In other words, discriminating well and poorly understood learning items had no influence on learner's expectations of information exchange concerning each of the learning items.

Consequently, Study 2 will try to replicate and verify if no relationship exists between monitoring and controlling. Moreover, the study wants to reveal if there is also no reverse relationship between controlling and monitoring. While in chapter 2 the study used a highly controlled learning situation with a fictitious learning partner, in Chapter 3 a similar controlled study will be presented, but for which learners were grouped into real dyads. In this way, the interplay of learners' metacognitive processes, learners' interaction behavior and their learning gains could be studied in a whole cooperative learning cycle with an individual and a collaborative learning session. As the potential for knowledge gains were highest in a heterogeneous grouping, the study focused on dyads with different knowledge levels, which allowed me to analyze the interplay between cognitive and social processes from both low- and high-ability learners' perspectives.

Chapter 3

Study 2 – Social metacognition: Monitoring and controlling self and partner knowledge

Chapter is based on

Schubert, M., Buder, J., & Hesse, F. (submitted). Social metacognition and heterogeneity in pairs of learners: A study on monitoring and controlling self and partner knowledge. *Learning and Instruction*

3.1 Introduction

Over the last thirty years, small group learning has turned into an established pedagogical means to increase students' learning achievements. Through providing and receiving explanations learners foster own knowledge and can increase other group members' knowledge by filling their knowledge gaps in order to achieve their mutual goals. But how do small learning groups, and dyads as the most basic form of learning groups, regulate the flow of information? Do individual learners in a group represent their own knowledge to decide what information they can provide? And does it mean that they represent their partner's knowledge too? In research on individual learning, questions about the regulation of own knowledge representations have a long tradition, and they are typically addressed under the rubric of metacognition (Flavell, 1979; Nelson & Narens, 1994). However, while there is indication that learners actively reflect on the specific knowledge constellation within their group (Lou et al., 1996, 2001; Webb, 1991), comparatively few scholars started studying the regulative mechanisms of knowledge representations within small learning groups (for an overview of different frameworks see Hadwin et al., 2011; Winne et al., 2013). Therefore, this chapter focuses on the reciprocal influence of monitoring and controlling own and partner knowledge, and analyzes these metacognitive mechanisms within heterogeneous dyads.

Metacognition

Metacognition has always been concerned about how we perceive ourselves and is generally defined as 'thinking about thinking' or 'knowledge and cognition about one's own thoughts' (Flavell, 1979; Hacker, 1998, p.3; Nelson & Narens, 1994). As outlined in section 1.1, Nelson and Narens (1994) described structural and process components of metacognition. They proposed that a distinction can be made between an object-level and a meta-level of cognition. On the object-level, learners use various cues to estimate their current state of knowledge (i.e., self-representation). This self-representation can be *monitored* by transferring the state to the meta-level. On the meta-level, the learner holds a goal state, to which the self-representation can be compared. Depending on the discrepancy between self-representation and goal state, learners *control* their learning by transferring information from the meta-level to the object-level (i.e., by proposing a course of action). As both processes interconnect object-level and meta-level, it is suggested that they are positively interdependent from each other. Therefore, monitoring one's knowledge more accurately should lead to more accurate controlling (e.g., Nietfeld et al., 2005; Thiede et al., 2009; Winne & Perry, 2000). On the other hand, control accuracy (and subsequently having more cues about certain learning items) should lead to better monitoring accuracy (Koriat et al., 2006; Koriat, 2012).

Social Metacognition and Cooperative Learning in Small Groups

Jost et al. (1998) proposed that metacognition can be also found in the social context and reviewed many classical findings from social psychology in terms of social metacognition (see section 1.2). They concluded that humans also hold a representation about their social partners. Unfortunately, these partner-representations are generally biased and rely on persons' self-representation, a phenomenon that was called false consensus bias (Birch & Bloom, 2004; Marks & Miller, 1987). However, there is also evidence that partner-representations become more accurate the more learners communicate with each other, and the more they provide cues about their level of understanding (Kenny & West, 2010; Letzring et al., 2006; Reis et al., 2011). Moreover, learners feel more attracted and emotionally close to the partner, the more learners get to know each other, which in turn causes a higher willingness to help (Biesanz et al., 2007; Korchmaros & Kenny, 2006; Watson et al., 2000). Together, the results hint at the assumption that in case of a lack of partner-information, learners try to become acquainted with the partner, which provides them with more partner-cues (i.e. cue-based framework; Koriat, 1997) and allows them to judge their partners more accurately.

While monitoring partner-representations depends on different cues the learner has available (i.e. own knowledge and experiences, cues from the partner), the question arises what kind of activities or behaviors would exemplify controlling in social metacognition. Demonstrated in section 1.2, I assume that providing explanations and eliciting explanations through questions are the main control actions when it comes to knowledge exchange in order to achieve the group's task goals in small group or cooperative learning (Johnson et al., 2007; Lou et al., 1996, 2001; Webb, 1989, 1991). Suggested by literature on small group learning (see section 2.1) and by results of Study 1, heterogeneous groups are likely to show such control behavior. In this group constellation, high-ability learners mostly provide explanations and answers, whereas low-ability learners receive most of the help, which leads to a teacher-student role distribution within the group.

However, the impact of providing and receiving explanations on learner's knowledge gain is ambiguous. While on the group level, higher proportions of interactions in heterogeneous groups are associated with higher learning outcomes of the group (Gijlers & De Jong, 2005), Webb (1991), for example, reported that high-ability learners benefited more from learning with a low-ability partner than vice versa. On the other hand, Saleh et al. (2005) found that low-ability students learned more in heterogeneous groups than high-ability students. Similarly, Lou and colleagues (1996, 2001) found in their reviews that low-ability learners benefited more or at least equally well like high-ability learners from heterogeneous group constellations, which contrasts Webb's results. The effects on learning achievement are therefore unclear (e.g., van der Laan Smith & Spindle, 2007).

In sum, there is reason to assume that learners do not only monitor their own knowledge during small group learning, but also monitor their learning partners to an extent, and with some accuracy. Moreover, controlling in social metacognition should result in providing, eliciting, and receiving explanations during discourse. The best way to trigger these social metacognitive behaviors is likely to be in heterogeneous dyads, but how elaborated explanations impact self- and partner-representations once they are provided or received is an open question. To shed light on this question, I refer to the developed extension of Nelson and Narens' (1994) object-metal-level model to dyadic interaction.

MEDIA: Metacognitive model in dyadic interaction

MEDIA has been developed with the goal to explain how learners represent own and partner knowledge, how these representations are monitored and influence control behavior within the dyad, and how provided and received information influence the monitoring of self- and partner-representations, as well as learning. In section 1.3 (see also Figure 1), I argued that controlling the exchange of elaborated explanations accurately depends on accurate monitoring of self- and partner-representations of knowledge. More specifically, MEDIA suggests that more explanations are provided for learning items for which the learner holds more knowledge compared to the partner (i.e. positive intragroup difference) and that more explanations are elicited for learning items for which the learner holds less knowledge compared to the partner (i.e. negative intragroup difference). The better learners assess their own knowledge, the better help-giving (i.e. providing explanations) and help-receiving (i.e. eliciting explanations) actions should fit a learner's knowledge needs. Study 1 (see section 2.5) showed that learners evaluate their own and their partner's individual discrepancies by adjusting the total amount of help-giving and help-receiving to either own or partner knowledge discrepancies. Furthermore, learners showed that they take intragroup differences of each learning item into account, as relative control accuracy measures indicated small proportions of accuracy. However, no adjustments could be found in regard to learner's relative monitoring accuracy. That means, better assessing well and poorly understood learning items had no influence on learner's control decisions of providing/receiving more or less information. To see if this effect persists in real dyads is one of the research questions in this chapter.

Moreover, MEDIA assumes that provided and received information helps learners to maintain their self- and partner-representations. Particularly, providing questions and explanations enables learners to monitor their self-representation more accurately (and their partner to gain knowledge when an explanation was provided), while receiving questions and explanations enables learners to monitor their partner-representation more accurately (and to gain knowledge when they received an explanation). Whether this relationship between control behavior and monitoring accuracy exists is another research question in this chapter.

The present study

The two research questions stated above have been translated to four concrete assumptions, tested in an empirical study involving heterogeneous dyads. First, the empirical study on heterogeneous dyads wants to replicate results on individual discrepancies of Study 1 and assesses in absolute figures whether high-ability learners actually provide more explanations, and whether low-ability learners will ask more questions. Second, it will be analyzed whether the no-relationship-effect between learner's own monitoring accuracy and their control accuracy persists in real dyads and assesses therefore monitoring accuracy (prior to the collaborative learning session), as well as control accuracy to provide or elicit explanations (during the collaborative learning session). For instance, high-ability learners in heterogeneous dyads should provide more explanations for learning items relative to other learning items (relative control accuracy) if and only if they can accurately assess that they *have* more knowledge of this learning items compared to other learning items (relative monitoring accuracy). While the first two questions refer to the relation between monitoring and controlling, the third question investigates the opposite direction: here it will be tested whether cues provided and received in dyadic interaction will predict the accuracy of self-representation and partner-representation. And finally, it will be tested whether relative group control accuracy has an impact on learning gains. To provide answers to these questions, I set up a classic cooperative scenario, consisting of an individual learning phase (i.e., reading comprehension), and a collaborative learning phase. I made sure that after the individual learning phase dyads perceived themselves as a heterogeneous group and took over either the role of a high- or a low-ability learner. Through coding and analysis of the collaboration, I could differentiate between help-receiving actions (i.e., asking questions) and help-giving actions (i.e., comprehension statements and explanations).

In heterogeneous groups, learners take over different roles, thus focusing on different learning items. I assume that due to their student role low-ability learners have a higher attention to learning items for which they can receive information. In contrast, high-ability learners are assumed to have a higher attention on learning items for which they can provide information, as they take the teacher role. In particular, I hypothesize that:

H1: Participants who judge themselves as the high-ability learner in a heterogeneous dyad will provide most of the explanations during a collaborative learning phase, while participants who judge themselves as the low-ability learner will ask most of the questions.

From the literature in metacognition it is known how important discrimination abilities of well and poorly understood learning items are in order to select appropriate learning items for the next step (relative monitoring accuracy; (Dunlosky & Hertzog, 1998; Schraw, 2009a, 2009b; Thiede et al., 2003). But as stated above, low- and high-ability learners focus on different aspects of knowledge and therefore use questions and explanations differently to regulate their learning process.

H2a: Learners high in relative monitoring accuracy will attain a high relative control accuracy of asking questions, but only when they judge themselves as the low-ability learner in the heterogeneous dyad.

H2b: Learners high in relative monitoring accuracy will attain a high relative control accuracy of providing explanations, but only when they judge themselves as the high-ability learner in the heterogeneous dyad.

The third hypothesis addresses the relation between social interaction and monitoring accuracy. MEDIA suggests that self-representations of knowledge become more accurate through the amount of cues that learners provide themselves (the explanations they give, or the questions they ask). Conversely, partner-representations of knowledge should become more accurate through the amount of cues that the partners provide (the explanations they give, or the questions they ask), though the false consensus bias might suggest that partner-representations are not highly accurate to begin with.

H3a: The more learners provide cues for their partners, the more accurate learners are in self-monitoring themselves (i.e., absolute self-monitoring accuracy). This is exemplified by a stronger association between their self-judgment of knowledge and their actual performance, which is mediated by the amount of cues provided during the collaboration.

H3b: The more learners receive cues from their partners, the more accurate learners are in monitoring their partners (i.e., absolute partner-monitoring accuracy). This is exemplified by a stronger association between the judgment of the partner and partner's actual performance, and it is mediated by the amount of partner's cues during collaboration, while being controlled for learner's own knowledge representation.

As mentioned in section 1.1 and 2.1, monitoring accuracy should lead to better and more appropriate controlling behavior, and this in turn should lead to better learning performance. MEDIA suggests that in heterogeneous dyads, low-ability learners should learn more than high-ability learners. This is because there are two avenues towards learning gains. First, if low-ability learners are better able to ask questions (thereby exhibiting high relative control accuracy; H2a), there will be more opportunities to fill their knowledge gaps. Second, even if low-ability learners do not ask questions, they are still likely to benefit if their high-ability learning partners are able to accurately provide explanations (thus exhibiting high relative control accuracy; H2b).

H4a: Learners, who judge themselves as the low-ability learner in a heterogeneous dyad, will achieve a higher learning gain than high-ability learners.

H4b: Learners high in relative control accuracy to ask questions will have higher learning gains than learners low in relative control accuracy, which will show up more for low-ability learners than for high-ability learners.

H4c: Learners high in relative control accuracy to provide explanations will affect partner's learning gains more than learners low in control accuracy, which will show up more for high-ability learners than for low-ability learners.

3.2 Method

Participants and Design

One hundred fourteen students were recruited from a university in Germany and randomly assigned to 57 dyads. One member of a dyad was assigned to the high-ability condition for the first learning phase, the other was assigned to the low-ability condition. In this first learning phase, both learners had the initial task of individually reading text paragraphs. In order to create heterogeneity within the dyads, learners in the low-ability condition were distracted by an additional task, receiving a radio report to listen to, while reading the paragraphs. During the reading task, both learners had to indicate judgments of comprehension (JOC) for each paragraph. If the difference of average JOC between learners in the high-ability condition and the low-ability condition was lower than a critical value of 15 percentile points, dyads were excluded from further analysis. The critical threshold of 15 points was determined through a pre-study (N = 88) where 39.2 % of participants rated a difference of 10 percentile points as heterogeneous, and 93.2 % rated a difference of 20 percentile points as heterogeneous. Thus, 14 dyads were omitted from the analyses. From the final selection of participants, 74 % were female and aged between 19 and 29

(*M* = 22.38, *SD* = 2.52). The random assignment brought up 16 mixed-sex and 27 same-sex dyads.

Material

Twelve interrelated text paragraphs about "Romans and the Germanic Tribes" were extracted from a German history textbook (Hilsch, 2012), each containing 64 to 69 words, and four to five sentences. For each paragraph, I created a comprehension multiple-choice test item with five answer options. In a preliminary study with 24 students from the same cohort, participants read each paragraph approximately 35s on average (*M* = 35.93, *SD* = 9.83), so that reading was limited to that time frame. Moreover, text material was rated fairly well according to legibility (*M* = 4.55, *SD* = 0.81; 1 = *not legible at all*; 6= *very legible*) and understandability (*M* = 4.54, *SD* = 0.61; 1 = *not understandable at all*; 6 = *very understandable*).

For low-ability-learners I selected a 7 minute radio report from a national, public radio station (Deutschlandfunk Wissen), which was not related to the text material but should nonetheless be of interest for learners. Hence, a report about vitamins and minerals in daily food consumption was selected.

A computer-based, experimental environment was used for questionnaires, text learning, collaborative learning and performance tests. Computers, which played the radio report for learners in the low-ability condition, were set to a maximum sound volume of 60dB. For the collaborative learning phase, learners used a chat in a 35-line conversation-window without scroll functionality.

Measures

The first set of measures refers to variables that were captured prior to collaboration.

Pre-collaborative judgements of comprehension (pre-JOC). During the individual learning phase, participants had to read twelve paragraphs. After each paragraph, they were asked "How well will you be able to answer to comprehension questions concerning the paragraph, when you will be asked in 20 minutes?" and had to indicate an integer between 0 and 100 (0 = *very badly*; 100 = *very well;* in increments of 1). All twelve judgments were averaged and compared to the partner's average score. When participants were at least 15 percentile points below their partner, they were coded

-1 (low-ability), while the partner was coded +1 (high-ability). Pre-JOCs were needed to compute relative monitoring accuracy.

Pre-test performance. Before the collaborative phase, learners were taking a comprehension test involving 12 multiple-choice questions. Performance was computed as percentage score of correct answers. Pre-test performance scores were needed to compute relative monitoring accuracy.

Relative monitoring accuracy. Each participant's *relative monitoring accuracy* for the individual learning phase was computed using gamma correlation between a participant's pre-JOC and the corresponding pre-test item across all paragraphs (e.g., Rawson et al., 2000; Thiede et al., 2003; for a discussion about the use of gamma in metacomprehension see Gonzalez & Nelson, 1996). Gamma rank correlations range between -1 and +1, where -1 indicates very low accuracy (i.e., systematically judging poorly understood learning items very high and well understood learning items very low), and +1 very high accuracy (i.e., systematically judging poorly understood learning items very low and well understood learning items very high).

The next set of measures addresses variables that express the actual information exchange among learners.

Amount of explanations / *amount of questions*. Based on the RAINBOW scheme for categorizing collaborative discussions (Baker et al., 2007), a code book had been created including seven distinct categories. Using the codebook, three raters, unaware of manipulations and the experimental design, coded the 57 chat protocols, while 11 protocols were coded by all three raters reaching a Fleiss' kappa interrater-reliability of 0.637 (substantial agreement: Landis & Koch, 1977). For the analyses presented in this paper, only three RAINBOW categories were used: comprehension statements and explanations were lumped together as help-giving actions (for brevity named as "explanations"), whereas questions were taken as indicator of help-seeking behavior.

Relative control accuracy. In order to arrive at relative control accuracy, I first computed individual difference scores between a learner's pre-JOC values for each paragraph (a positive difference indicating that a learner knew more than her partner). For each paragraph the difference scores were Pearson-correlated with the number of explanations the learner provided (relative control accuracy for providing explanations), and Pearson-correlated with the number of questions the person has asked (relative control accuracy for asking questions). High relative control accuracy for providing explanations yields a positive value (the more a learner knows in comparison to the partner, the more explanations will be provided). Conversely, high relative accuracy for asking questions yields a negative value (the less a learner knows in comparison to the partner, the more questions the learner will ask).

The final set of measures addresses variables that were captured after dyadic collaboration.

Post-collaboration judgment of own comprehension (post-JoOC) / post-collaboration *judgment of partner comprehension* (post-JoPC). After the collaboration ended, I assessed each learner's general JOC concerning the whole material, and asked them additionally how they estimate their learning partner's general JOC. Therefore, I used similar questions like in the individual learning phase ("How well will you/your learning partner be able to answer to comprehension questions in a test in about 10 minutes?"), to which participants had to answer by entering an integer between 0 and 100 in a text field (0 = *very badly*, 100 = *very well*).

Post-test performance / learning gains. After collaboration, learners received the same multiple-choice questions as in the pre-test. Performance was computed as the percentage score of correct answers across all twelve test items in each test, while *learning gains* were computed by calculating percentage difference scores between learner's test results of pre-test and post-test.

Procedure

Following Thiede's approach, the experiment incorporated an individual learning phase with immediate JOCs, a pre-test, a collaborative learning phase to regulate their knowledge for each learning item and a final post-test (de Bruin et al., 2011; Thiede et al., 2003). At arrival, the two participants were first seated next to each other at a table, where a computer with a table microphone was placed. Right at the beginning, participants were informed that the experiment involved an individual learning phase where they studied text material, followed by a collaborative learning phase where they have the opportunity to discuss and exchange their knowledge to gain more insight. Further, they were made believe to defend their learning results at the end of the experiment in front of the experimenter to generally increase motivation in collaborating with the other person (i.e., impression management: Leary & Kowalski, 1990). After being briefed, participants were placed in separate rooms with a computer and headphones. The experiment started with a questionnaire on demographic data. Then, participants received their (interdependent) task goal (i.e., try to achieve 90 % on average with your partner in a final test after the collaboration), as well as instructions for the first, individual learning stage. Additionally to basic instructions, dual-task-learners were made believe that after the first learning phase they would have to complete a knowledge test about both text material and the radio report. Both learners were then asked to take their headphones on. In the individual learning phase, each of the twelve paragraphs appeared successively on separate screens with a reading time limit of 35s, but only dual-task-learners were presented with the radio report at the same time. Each text screen was followed by the corresponding pre-JOC question on another screen. After all twelve paragraphs were presented, a distracter text appeared for 30s to clear working memory. Participants then answered to the twelve test items of the pre-test, which had been shuffled into a random order beforehand and presented as single trials on the screen. Contrary to the announcement, low-ability learners did not have to answer questions related to the radio report. Once both learners finished the pre-test, a screen appeared showing the averaged pre-JOC values of both learners in a bar chart design. This was done to ensure that learners perceived themselves as a heterogeneous dyad. On continuing, learners entered the chat room for the second learning stage, without having their own or their partner's pre-JOC scores available. During the collaborative stage, learners had the possibility to interact for a maximum of 30 min, but could collectively decide to abort the chat any time. After learners left the chat screen, each of them had to indicate both post-JoOC and post-JoPC. After reading another distracter text, learners finally answered to the same twelve comprehension test items a second time, shuffled into a new order (post-test). After telling the participants that the originally announced post-collaboration interview would be canceled, the experiment finished by remunerating and debriefing participants.

Dyadic Data Analysis

Some of the hypotheses in this study are about relations among variables that originate from both learners separately (for instance, the prediction that learning gains of low-ability learners may depend on the controlling accuracy of their partners). However, such data are not independent from each other. To account for this, the actor-partner-interdependence-model (APIM: Cook & Kenny, 2005; Kenny et al., 2006) was used for the analysis of all hypotheses except H1. APIMs were estimated using general least square analysis with correlated errors and restricted maximum likelihood estimation (R-script "APIMM.R" provided by David Kenny and Thomas Ledermann [2]). In the APIM analyses, the learners were always treated as distinguishable. Consequently, the model allows predicting separate estimates for both learners. Effect sizes for low- and high-ability learners are given as partial correlations. To test whether the coefficients of low and high level learners are significantly different from each other, Z tests were calculated. Moreover, I always report the combined learner effect of predictors, which indicates an effect when learners were treated as indistinguishable.

3.3 Results

In Hypothesis 1, I assumed that low-ability learners will ask more questions during the collaborative learning phase than high-ability learners, while high-ability learners will provide more explanations than low-ability learners. In using within-dyad contrasts denoted in Table 5 the first two within contrasts represent the hypothesis stated above, while the third one represents a residual contrast capturing the remaining variance.

[2] Retrieved the 16th of June 2015; By date of submission, an updated web-based app permits the computation on https://davidakenny.shinyapps.io/APIM_MM/ , which reveals the same results like the R-Script

Table 5. Within-dyad contrasts (WC) and the residual contrast (RC) representing Hypothesis 1.

	Categories of utterances				
	Questions		Explanations		
	Low-ability	High-ability	Low-ability	High-ability	
WC1	+1	-1	0	0	Low-ability learners ask more questions
WC2	0	0	-1	+1	High-ability learners provide more explanations
RC	-1	-1	+1	+1	Generally, dyads exchange more explanations than questions

While the first within contrast showed no difference from zero (M = 0.30, SD = 0.54, t(42) = 0.564, p = .576, r = .087), the second within contrast revealed that high-ability learners provided significantly more explanations than low-ability learners (M = 3.63, SD = 1.17, t(42) = 3.093, p = .004, r = .431). Moreover, learners also generally exchanged more explanations than questions, as the residual contrast demonstrates (M = 9.37, SD = 1.21, t(42) = 7.775, r < .0001, r = .768). Thus, results indicate that learners do not use much questions during knowledge exchange, but that high-ability learners reacted on the perceived intragroup difference by providing more explanations than low-ability learners.

Generally, Hypothesis 2 was addressing the relation between relative monitoring accuracy (pre-collaboration) and relative control accuracy (during collaboration). More specifically, H2a predicted that low-ability learners will exhibit a stronger association between monitoring accuracy and control accuracy for asking questions than high-ability learners. In contrast, H2b predicted that high-ability learners will exhibit a stronger association between monitoring accuracy and control accuracy for providing explanations than low-ability learners. Thus, two separate APIMs are computed—one for each sub-hypothesis—of which the actor results are presented in Table 6.

Table 6. Actor effects of separate APIMs of relative monitoring accuracy (RMA) for predicting relative control accuracy for asking questions (RCA-Q; H2a) and for predicting relative control accuracy for providing explanations (RCA- E; H2b).
Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and effect size (r). The test of main differences between low and high-ability learners (intercept) is indicated by a p-value, while moderating effects of predictors are Z-test, also including p-values.

		Δ R^2	B	SE	*β*	p	r	p/ Z(p)
RCA-Q								
Intercept	Low-ability		-0.03	0.06		.636		.499
	High-ability		0.03	0.06		.614		
	combined		0.00	0.04		.994		
RMA	Low-ability		0.08	0.09	0.13	.402	.132	-1.985
(actor	High-ability		-0.18	0.09	-0.29	.052	-.298	(.051)
effect)	combined		-0.05	0.07	-0.08	.437	-.087	
Model	Low-ability	.102						
	High-ability	.043						
RCA- E								
Intercept	Low-ability		-0.02	0.06		.704		.614
	High-ability		0.02	0.07		.779		
	combined		-0.00	0.05		.956		
RMA	Low-ability		-0.01	0.10	-0.02	.898	-.020	2.305
(actor	High-ability		0.34*	0.11	0.45	.003	.432	(.024)
effect)	combined		0.16*	0.08	0.22	.036	.232	
Model	Low-ability	.136						
	High-ability	.168						

With regard to H2a, the Z-test for the actor effect failed to reach statistical significance. In other words, there was no difference between high-ability and low-ability learners with regard to their relation between monitoring accuracy and control accuracy for asking questions. As high accuracy is associated with negative values, further inspection of the data suggests that it was high-ability learners who showed a somewhat stronger (but only marginally significant) relationship between monitoring accuracy and control accuracy for asking questions. Thus, H2a has to be rejected.

In contrast, H2b could be confirmed. As expected, high-ability learners with a high monitoring accuracy also exhibited higher control accuracy for providing information. In other words, only when high-ability learners showed strong abilities to discriminate between well and poorly understood learning items, they were also better in evaluating the difference between own knowledge and partner knowledge, and adjusted their explanation-giving accordingly. Conversely, this relation was not found for low-ability learners, resulting in the expected significant difference among the ability conditions.

Next, H3a assumes that learners are more accurate in predicting their own knowledge after the collaborative learning phase when they provide a high number of cues (i.a. absolute self-monitoring accuracy). This should be captured by a strong association between self-judgment (post-JoOC) and their actual performance and is mediated by the information they provided (explanations and questions). Even though no difference in provided cues between low- and high-ability learners was hypothesized, I analyzed self-monitoring accuracy for possible ability differences as H1 revealed that high-ability learners provided more explanations than low-ability learners and also generally more explanations were given than questions asked. Similar, I checked for differences between low- and high-ability learners in partner-monitoring accuracy in H3b, though no differences in received cues were originally expected. In H3b I argued that learners have a higher partner-monitoring accuracy, represented by a stronger association between their partner judgment (post-JoPC) and their partner's actual performance, which is mediated by increased information reception (questions and explanations) from their partner. Because high-ability learners provided more cues, the mediation effect might be stronger for low-ability learners. Due to false consensus, I controlled partner-representations by judgements of own comprehension (post-JoOC).

In order to test Hypothesis 3, the truth-and-bias-model by West and Kenny (2011) was applied as the principle paradigm (Figure 2 and Figure 3). The general idea of the paradigm is that the judgement of comprehension (post-JoOC for H3a, post-JoPC for H3b) is treated as the outcome variable while the actual post-test performance value (self or partner) is treated as the predictor, called truth variable. Moreover, the truth-and-bias model permits to add one or more supplementary predictors (called bias variables), which can take the role of mediators or control variables. H3a is represented by Figure 2, in which learner's actual performance is the truth variable predicting learner's post-JoOC, which in turn is potentially being mediated by the amount of expressed cues (questions and explanations). H3b is represented by Figure 3, in which partner's actual performance is the truth variable predicting learner's post-JoPC, potentially mediated by partner's amount of expressed cues and controlled for learner's post-JoOC.

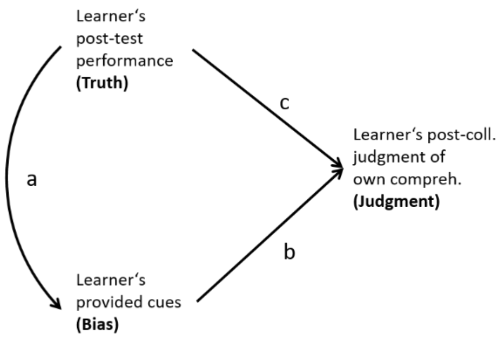

Figure 2. Model of absolute *self*-monitoring accuracy in Study 2
Model tests absolute self-monitoring accuracy with provided cues as one bias variable (West & Kenny, 2011)

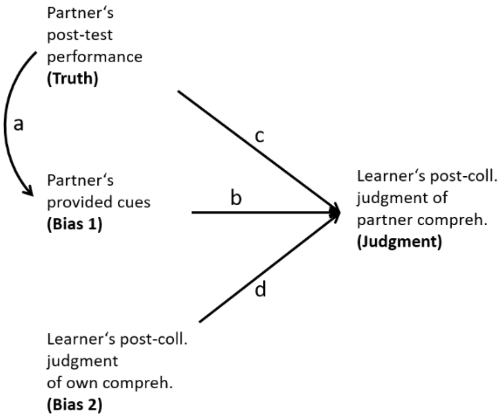

Figure 3. Model of absolute *partner*-monitoring accuracy in Study 2
Model tests absolute partner-monitoring accuracy with received cues and judgement of own comprehension as two separate bias variables (West & Kenny, 2011)

I followed West and Kenny's (2011) recommendation for centering the variables in a way that the intercept of the model directly indicates whether learners over- or underestimate themselves or the partner. Therefore, in both models displayed in Figure 2 and 3 all variables are centered by the grand mean of the truth variable except for the cues, which were centered by their own grand mean. Using this centering strategy, overestimation is indicated by a significant positive intercept, while underestimation is indicated by a significant negative intercept. To calculate all paths, two APIMs for each model in Figure 2 and 3 needed to be estimated. Results regarding H3a are presented in Table 7.

Table 7. Effects of two separate APIMs. The first APIM tested actor effects of learner's post-JoOC predicted by learner's post-test performance and by learner's provided cues, while the second APIM tested actor effects of learner's post-test performance on learner's provided cues (questions and explanations).
Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and effect size (r). The test of main differences between low and high-ability learners (intercept) is indicated by a p-value, while moderating effects of predictors are Z-test, also including p-values.

		Δ R^2	B	SE	*β*	p	r	p/ Z(p)
post-JoOC								
Intercept	Low-ability		-30.48	3.15		<.001		<.001
	High-ability		-16.84	2.67		<.001		
	combined		-23.66	2.43		<.001		
c (actor effect)	Low-ability		0.47*	0.17	0.36	.009	.339	0.469
	High-ability		0.60*	0.21	0.46	.005	.421	(.641)
	combined		0.54*	0.13	0.38	<.001	.399	
b (actor effect)	Low-ability		0.59	0.56	0.22	.298	.167	-0.017
	High-ability		0.58	0.37	0.22	.120	.247	(.986)
	combined		0.59	0.30	0.20	.053	.199	
Model	Low-ability	.264						
	High-ability	.367						
Cues								
Intercept	Low-ability		-0.65	1.06		.542		.036
	High-ability		1.71	1.38		.220		
	combined		0.53	1.10		.633		
a (actor effect)	Low-ability		0.14*	0.06	0.29	.017	.360	-0.547
	High-ability		0.07	0.11	0.14	.534	.098	(.586)
	combined		0.10	0.06	0.22	.071	.194	
Model	Low-ability	.092						
	High-ability	.001						

The intercept of the first APIM indicates that both high-ability learners and low-ability learners underestimated their performance. However, the significant main effect shows that high-ability learners' self-judgments were more accurate (i.e., closer to zero). H3a predicted that this effect can be explained by the amount of provided cues, which was not the case. First, there is a strong relation between judgements of comprehension and actual performance (path c), and this relation occurred for both types of learners, and not only for high-ability learners. And second, the non-significant associations between provided cues and judgements of comprehension (path b) and between actual performance and provided cues (path a), do not suggest any mediation effects by provided cues. Additionally, no differences were found between low- and high-ability learners, as the Z-test scores were non-significant. In sum, high-ability learners are more accurate and underestimate themselves to a lesser extent, but this effect cannot be explained by the amount of provided cues. Table 8 shows the results regarding H3b, testing the model displayed in Figure 3.

Table 8. Effects of two separate APIMs. The first APIM tested effects of learner's post-JoPC predicted by partner's post-test performance (partner effect), partner's provided cues (partner effect) and by learner's post-JoOC, while the second APIM tested actor effects of partner's post-test performance on partner's provided cues (questions and explanations).
Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and effect size (r). The test of main differences between low and high-ability learners (intercept) is indicated by a p-value, while moderating effects of predictors are Z-test, also including p-values.

		Δ R^2	B	SE	*β*	p	r	p/ Z(p)
post-JoPC								
Intercept	Low-ability		1.29	3.61		.722		.944
	High-ability		0.95	3.58		.790		
	combined		1.12	2.70		.679		
c	Low-ability		0.16	0.17	0.14	.359	.152	-0.319
(partner	High-ability		0.09	0.11	0.08	.423	.133	(.751)
effect)	combined		0.13	0.10	0.11	.222	.142	
b	Low-ability		0.66*	0.27	0.29	.016	.379	-0.671
(partner	High-ability		0.36	0.34	0.16	.294	.174	(.505)
effect)	combined		0.51*	0.21	0.22	.017	.267	
d	Low-ability		0.54*	0.11	0.61	<.001	.647	0.791
(actor	High-ability		0.67*	0.12	0.76	<.001	.670	(.431)
effect)	combined		0.60*	0.08	0.76	<.001	.658	
Model	Low-ability	.542						
	High-ability	.689						
partner cues								
Intercept	Low-ability		1.71	1.38		.220		.036
	High-ability		-0.65	1.06		.542		
	combined		0.53	1.10		.633		
a	Low-ability		0.07	0.11	0.14	.534	.098	0.483
(actor	High-ability		0.14*	0.06	0.29	.017	.360	(.631)
effect)	combined		0.10	0.06	0.22	.071	.194	
Model	Low-ability	.001						
	High-ability	.092						

The intercept for both learners were non-significant thus showing that both learners are accurate in predicting their partner's performance. Further, the combined estimates for path c (relation between partner judgments and actual partner performance) and path b (partner's provided cues) lend support for H3b, as only path b became significant. In other words, learners estimated their partner's knowledge the higher, the more their partner provided explanations. However, as the relation between partner performance and partner cues (combined effect of path a) is not significant, no mediation could be established. Even though partner cues became a stronger predictor for low-ability learners than for high-ability learners, the mediation could not be established for low-ability learners alone, as the Z-test scores for path b and path a showed no significance in favor of low-ability learners. In other words, both learners in heterogeneous dyads are accurate in monitoring their partner, but this could not be explained by the amount of received information from the partner. Nonetheless, learners are influenced by what the partner provided, as well as by their own self-judgments (path d), which also became a strong predictor (combined effect) without a significant difference between low- and high-ability learners (non-significant Z-test score).

According to Hypothesis 4, low-ability learners should have higher learning gains than high-ability learners (H4a), particularly if they had better control accuracy for asking questions (H4b), and if their partners had better control accuracy for providing explanations (H4c). The assumptions were tested using one APIM with relative control accuracy measurements as predictor variables and learning gains as the outcome variable. Results are presented in Table 9.

Table 9. Effects of an APIM predicting learning gains by learner's relative control accuracy to ask questions (RCA-Q) and partner's control accuracy for providing explanations (RCA- E). Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and effect size (r). The test of main differences between low and high-ability learners (intercept) is indicated by a p-value, while moderating effects of predictors are Z-test, also including p-values.

		Δ R²	B	SE	*β*	p	r	p/ Z(p)
Intercept	Low-ability		6.55	2.16		.003		.016
	High-ability		0.48	1.15		.680		
	combined		3.51	1.22		.005		
RCA-Q (actor effect)	Low-ability		3.19	8.74	0.09	.716	.059	-0.913
	High-ability		-6.37	5.74	-0.17	.439	-.177	(.364)
	combined		-1.59	5.22	-0.04	.762	-.033	
RCA-E (partner effect)	Low-ability		3.18	7.40	0.10	.699	.069	-0.922
	High-ability		-4.98	4.86	-0.16	.309	-.164	(.360)
	combined		-0.90	4.43	-0.03	.840	-.022	
Model	Low-ability	.000						
	High-ability	.000						

The positive and significant intercept shows learning gains for low-ability learners, but not for high-ability learners, which is confirmed by the corresponding main effect score. Thus, H4a can be accepted. However, H4b and H4c need to be rejected, as the expected learning gains could not be predicted by a better control accuracy of low-ability learners asking questions or by a better control accuracy of their high-ability partners providing explanations. Having said that, it is unclear what caused low-ability learners to learn more during the collaboration than high-ability learners. As the efficacy of knowledge exchange seems to play no role, it is worth verifying whether the pure amount of questions and explanations are responsible for low-ability learners' knowledge gains. Thus, I ran a second APIM looking for the actor effect of (grand mean centered) questions on low-ability learners' knowledge gains, as well as on the partner effect of (grand mean centered) explanations on their learning gains. The results are presented in Table 10.

Table 10. Effects of an APIM predicting learning gains by learner's questions and received explanations from the partner.
Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and effect size (r). The test of main differences between low and high-ability learners (intercept) is indicated by a p-value, while moderating effects of predictors are Z-test, also including p-values.

		Δ R^2	B	SE	*β*	p	r	p/Z(p)
Intercept	Low-ability		5.17	2.26		.025		.138
	High-ability		1.44	1.25		.252		
	combined		3.31	1.33		.015		
questions (actor effect)	Low-ability		0.14	1.01	0.04	.887	.023	-0.741
	High-ability		-0.70	0.51	-0.18	.179	-.215	(.461)
	combined		-0.28		-0.07	.631	-.057	
expla-nations (partner effect)	Low-ability		0.65	0.33	0.39	.052	.305	-0.600
	High-ability		0.37	0.32	0.22	.256	.183	(.551)
	combined		0.51*		0.31	.030	.221	
Model	Low-ability	.053						
	High-ability	.007						

Again, the intercepts of both types of learner reveal that low-ability learners seem to learn more than high-ability learners, but this time the main effect was non-significant. Thus, both learners gain knowledge, when they ask the average amount of questions, and receive the average amount of explanations during the collaborative phase. However, the results indicate that mainly explanations of the partner are responsible for increased learning gains as the association is positive and significant when both learners are taken together. Even though, no significant differences have been found between low and high-ability learners, the standardized coefficients and effect sizes show that receiving explanations from high-ability learners seems more important to low-ability learners, than vice versa. With regard to asking questions, no relationship was found with learning gains.

3.4 Discussion

The present chapter describes an extension of Nelson and Narens' (1994) object-meta-level model to dyadic interaction (MEDIA), and it proposes an approach to study metacognitive activities of learners in a dyadic learning scenario. Basic assumptions of the model hold that better monitoring accuracy will lead to more efficient control behavior. In turn, higher control accuracy causes higher learning achievements, while the amount of information exchanged will lead to better monitoring accuracy, thus closing the loop between object-level and meta-level.

Focusing first on behavioral results, an increased total amount of help-giving actions was observed from high-ability learners, which correspond to results of chapter 2 and which were also reported elsewhere (Saleh et al., 2005; Webb, 1991). Additionally, high-ability learners were better able to transfer their higher monitoring accuracy to higher control accuracy of giving helpful information. In other words, if high-ability learners were able to discriminate among poorly and well understood learning items, they were providing explanations only about those items for which they possessed a higher comprehension level than their peers. For this, recognizing intragroup differences was essential. While the result supports the hypothesis of a relationship between monitoring and control accuracy, it also seems to confirm the hypothesis that high-ability learners focus more on help-giving actions than on help-receiving actions. In contrast, low-ability learners were not more focused on help-receiving actions than high-ability learners and not very effective in asking questions, neither with high nor with low monitoring accuracy. One potential reason for this might be the fact that only clearly formulated questions were coded as help-receiving actions. This strict coding scheme could have biased the results, and a qualitative discourse analysis should uncover more subtle help-receiving actions, which provoked further inquiries by the partner.

Even though high-ability learners achieved to provide help effectively, I did not find evidence that this efficient information transfer led to low-ability learners' knowledge gains, as it was found in metacognition of individuals (Thiede et al., 2003). In fact, a supplementary analysis revealed that rather the pure amount of received explanations contributed to low-ability learners' knowledge gains than their partners' control efficacy of providing explanations. However, I have to point out that generally both learners showed rather low abilities of relative monitoring accuracy.

Thus, helping learners to increase their relative monitoring accuracy, for example through delayed JOCs (Rhodes & Tauber, 2011) or through letting them generating their knowledge before judging their knowledge (Dunlosky & Lipko, 2007; Thiede et al., 2003), might lead to higher control accuracy, which in turn might have a higher impact on knowledge gains.

Generally, low-ability learners' knowledge gains during the collaboration supports findings from Lou et al. (1996). But at the same time, the relative lack of knowledge gains for high-ability learners is contradictory to findings of both Webb (1991) and Lou et al. (2001). However, high-ability learners might have already reached the performance ceiling after the individual phase with less potential to increase their knowledge during collaboration, which gives a first hint that the learning task was rather easy. As low-ability learners benefited more from the total amount of received cues than from the efficiency of received cues, they might have taken the advantage of being grouped with a high-ability learner to also double-check their understanding of learning items, which they indicated as high after the individual learning phase. Moreover, being efficient during collaboration was not the primary goal for the dyad. Therefore, more effort has to be put into identifying the exact circumstances of collaboration under which relative control accuracy is beneficial for learning achievement (e.g., time pressure; Kelly & Karau, 1999; or cognitive load; Sweller, 1994).

The study in the current chapter also investigated the relation between provided/ received explanations and post-collaborative self- and partner-monitoring accuracy. First, I found that both type of learners underestimated themselves to a high extent. However, high-ability learners were generally less biased. Results of the predictors indicated that the more learners knew (revealed through the comprehension test), the higher they rated themselves, but this effect could not be explained by the amount of provided cues. Together, the predictors explained only 26 % to 37 % of the variance. Therefore, other factors might have caused their underestimation such as a rather low task difficulty. Moore and Healy (2008) reviewed several studies investigating self-judgments after easy tasks, stating that "when people have imperfect knowledge of their own performance, the error in their estimates will make those estimates regressive […] when it (performance) is high". As performance rates reached on average around 79 % and learners had no direct feedback about their potential knowledge, they might have regressed their estimates to a, in their opinion, more plausible value of around 60-65 %—a value, which is often influenced by their prior experience with test type and performances with regard to the subject.

Second, learners were very accurate in estimating their partner's knowledge under conditions of average partner performance, average self-judgements of comprehension, and average amount of received information. These three predictors explained between 54 % and 69 % of the variance in partner judgments. By controlling for the false consensus bias, both type of learners rated their partners the better, the more cues partners provided. However, partner accuracy could not be explained only by how much knowledge the partner achieved to externalize; the false consensus bias itself became also a very strong predictor as learners tended to align their partner judgments with their self-judgments. Consequently, both factors served as proxies for estimating partner knowledge.

Together, the results underscore the relevance of MEDIA. I could demonstrate that the construction of self and partner-representations exists. Moreover, learners are capable of monitoring themselves and the partner, and they are capable of controlling for individual discrepancies. Depending on their perceived role in the dyad, learners provided more or less explanations. Similar to Study 1, this study provided evidence that learners take intragroup differences into account, as high-ability learners' relative control accuracy of providing explanations became significantly different from zero when relative monitoring accuracy was high. In order to get more efficient in providing explanations, learners needed to evaluate intragroup differences. Even though partner control accuracy of providing explanations had no influence on learners' knowledge gains, it can be stated that the total amount of received explanations updated learners' knowledge states and led to learning. Moreover, partner-judgments were influenced by the amount of received cues. The more cues the partner provided, the higher her knowledge was rated by the learner. Thus, receiving cues seems to have the expected functionality presented in MEDIA in updating the learner's own knowledge and informing about their partner's knowledge state. In contrast, the amount of provided cues did not foster a learner's self-representation, as it was expected in MEDIA. In sum, several parts of the model were elicited through the experimental setting, but other parts such as the influence of provided cues, investigations of individual discrepancies with different goal states, as well as the effect of control accuracy on learning need further examination.

The results provide exciting insights into how cognitive and intragroup processes are intertwined. However, some limitations with regard to generalizability should be mentioned. First, the cooperative learning task that I used is not generalizable to all group learning scenarios. Thus, it might also be interesting to examine purely collaborative or, in contrast, more structured methods such as jigsaw (Aronson, 1997), in which not ability differences but distributed and complementary abilities can be investigated. Second, the learning material about history appeared to be rather easy, so one should be careful to draw conclusions about other topics (e.g., logical thinking in math or science). And finally, learners collaborated in a chat environment, in which certain context information (e.g., mimic, gestures, or emotions) gets lost. Natural face-to-face conversations might help learners to adapt even more to their partners' knowledge needs.

The experimental design captured self and partner judgments of knowledge only at the end of the collaboration and with regard to the learning material overall. In future studies, capturing intermediate judgements of comprehension over the course of information exchange could help to identify when cognitive biases or process variables become more or less influencing. Moreover, judging each paragraph separately for self and partner knowledge might better show how relative monitoring accuracy and relative control accuracy change over time. Findings about metacognitive regulation in groups could provide a basis to improve skill training in group processes (e.g., Nam & Zellner, 2011), perspective taking skills (e.g., Gockel & Brauner, 2013), or subtle knowledge awareness tools (Dehler et al., 2011; Engelmann & Hesse, 2010; Sangin et al., 2011).

3.5 Interlude

In chapter two and three, results showed that dyads, heterogeneous in their levels of understanding of the to-be-learned material, have a high potential of exchanging information. More specifically, both high- and low ability learners seem to hold own and partner knowledge states in mind when interacting with each other. During the collaborative learning session, high-ability learners have a stronger tendency to provide explanations, while low-ability learners expect and try to elicit explanations.

Therefore, perceiving certain levels of knowledge triggers respective interactions. Moreover, received explanations help learners to gain knowledge and to adjust their representations of partner knowledge. While Study 1 and Study 2 lend some support to the notion that self and partner representations play a pivotal role in the context of (mutually) explaining learning content, one should be careful to generalize these finding to other collaborative activities. For instance, it is unclear whether learners in a cooperative learning task also represent own and partner states of knowledge in the realm of a problem solving task. I would argue that in problem solving tasks monitoring the states of individual prior knowledge and recognizing intragroup differences are equally important cognitive processes that are needed to successfully exchange information and solve a problem. The focus of the following study is to test whether learners holding complementary knowledge prior to a collaborative problem-solving session will engage in more interactions than learners holding identical knowledge. Interactions in a problem solving task are not only meant to provide and elicit explanations of an existing knowledge, but also to create or bring up new ideas and solutions, thus new knowledge. If learners with complementary prior knowledge interact more with each other, they should in turn be more accurate in monitoring own and partner knowledge of the solution generated during the task. Therefore, the study in the next chapter will further examine the range of validity of MEDIA and test the model in a more authentic cooperative problem solving task including a computer-supported face-to-face collaboration.

Chapter 4
Study 3 – Positive resource interdependence in the context of scientific and cooperative information problem solving

Chapter is based on

Schubert, M., Buder, J., & Hesse, F. (submitted). Positive resource interdependence on promotive interactions, group performance and individual free recall in the context of scientific and cooperative information problem solving. *Journal of Applied Cognitive Psychology*

4.1 Introduction

Searching and finding information, whether in the professional or educational context, is one of the most important skills in the 21st century and becomes even more important in the realm of ubiquitously accessible information through the World Wide Web. Searching and extracting information from different sources have been labeled under the rubric of “information seeking” or “information retrieval” as part of information literacy (Ingwersen, 1996; Kulthau, 1991; Marchionini, 1995). From an educational perspective, searching for information can be seen as an information problem, through which not only domain knowledge, but also a variety of skills is acquired (information problem solving [IPS]; Brand-Gruwel, Wopereis, & Walraven, 2009). However, as tasks in IPS are often authentic in nature, and thus complex and ill-defined, learners need substantial support to engage into all required cognitive processes (Walraven, Brand-Gruwel, & Boshuizen, 2008; Wopereis, Brand-Gruwel, & Vermetten, 2008). Instead of focusing on instructional methods for individual IPS (Gerjets & Hellenthal-Schorr, 2008; Stadtler & Bromme, 2008; Walraven et al., 2008), this paper attempts to incorporate cooperative methods for IPS tasks and wants to replicate effects of positive resource interdependence (Bertucci, Johnson, Johnson, & Conte, 2011; Johnson, Johnson, & Stanne, 1989; Ortiz et al., 1996) within dyads solving an information problem.

A quasi-experiment is carried out to test whether positive resource interdependence (i.e., learners holding complementary knowledge) affects the joint search process of scientific information. Together with positive outcome interdependence (i.e., shared goals), positive resource interdependence is claimed to increase promotive interactions within the learning group—that is helping and encouraging the learning partner, but also creating ideas and generating solutions by the exchange of information (Bertucci et al., 2011; Johnson et al., 1989; Ortiz et al., 1996). Therefore, the research question is pursued in two ways: on the one hand by exploring whether promotive interactions have an impact on groups’ actual search actions; and on the other hand by testing whether promotive interactions affect learners’ recall abilities of compiled items after the search task. To get more insight into patterns of knowledge representations, and to see whether promotive interactions about the search process influence own and partner representations of recall abilities, the self-other-agreement approach from social perception will be introduced to explore learners' prediction accuracy of own and partner recall abilities (Kenny & West, 2010).

Information Problem Solving

A major skill that students have to achieve during their educational life is to correctly search and filter external information to solve their tasks, whether they retrieve information from books or hypermedia environments. In fact, searching for information has been described as part of the typical information problem-solving process (Brand-Gruwel et al., 2009; Kulthau, 1991). Brand-Gruwel et al. (Figure 1; 2009) described how the process of solving an information problem on the Internet might unfold: a) defining the information problem, b) search information, c) scan information, d) process information and e) organize and present information. All these steps are constituted by different (sub-)skills. For instance, searching for information requires learners to derive search terms and to evaluate search results, while processing information requires learners to elaborate on the content and evaluate the information (Wopereis et al., 2008). Hence, solving information problems requires learners to perform a variety of cognitive activities. Moreover, it is argued that searching for information in hypermedia environments (like the Internet) is a loosely structured process as the strategies to find the information, the selection of search terms and the way searchers elaborate on the results can lead to multiple paths through the information source (Jonassen, 2000). Such problem-solving tasks, for which just a low structure exists, are often defined as ill-structured. An example, which is used in the present study, is searching for relevant keywords and authors in regard to an initially given research topic—a situation which often occurs for scientists or for students when they explore a new research field. Even if the starting point and goal state of a search activity might be clearly defined, each person might use different strategies (e.g., concentrating on authors vs. concentrating on keywords), use different search terms, decide to take different rules and criteria for judging information, and therefore process and extract different information from the content. Consequently, every information seeker might come up with a different set of keywords and authors at the end of the task.

It has been discussed among educators whether learners should be confronted with such ill-structured tasks in totality, or whether tasks should be divided into sub-tasks for each to-be-learned sub-skill. Van Merriënboer and Kester (2008) have promoted the former instructional approach, and they proposed the use of *authentic tasks* that require learners to engage with all the different aspects, skills and processes. Authentic tasks are thus aimed to motivate students to apply and integrate every cognitive skill, and to emphasize the overall task performance of learners.

However, merely implementing IPS into an authentic scenario of real-life problems might not suffice. As Brand-Gruwel and colleagues underscore (Walraven et al., 2008; Wopereis et al., 2008), learners need further instructional support in order to learn and benefit from search environments. For example, special training programs to improve learners' Internet search skills showed positive effects on learners' declarative knowledge and search performance (Gerjets & Hellenthal-Schorr, 2008). Moreover, Stadtler and Bromme (2008) showed that receiving evaluation prompts during the search for medical information improved laypersons' knowledge about different resources, while monitoring prompts improved their acquisition of facts (even though comprehension was not improved). Another way to help learners engaging more with the search process and to elaborate more on the found information, is through collaboration among learners (Walraven et al., 2008). However, research on collaborative IPS tasks is scarce (Hyldegard, 2006).

Just recently, researchers started investigating the effectiveness of groups in finding information (Shah, 2014). Lazonder (2005) compared individual and dyadic information seekers, discovering that pairs are not only faster in retrieving information from the Internet, but also exhibit a richer repertoire of search terms. Similar results have been reported by Dinet and Vivian (2012) studying grade three pupils. But to my knowledge, no further research exists that studies search efficiency in groups, or different collaborative scripts/scaffolds during mutual search activities. Therefore, the chapter wants to make a first step towards studying cooperative learning with IPS tasks. Adapted from the field of cooperative learning, I was particularly interested in the question of whether positive resource interdependence, that is providing learners with different but complementary material prior to their collaboration, might improve learners' search and problem solving behavior.

Positive interdependence among information seekers

Fifty years of cooperative learning research has demonstrated that the mechanism of cooperation, in contrast to competition and individualistic learning, has a facilitating effect on learners' knowledge gains (Johnson et al., 2007). One of the basic prerequisites of cooperative learning is that learners need to feel positively interdependent from each other (Johnson & Johnson, 1999). Only when learners feel interdependent they will engage in promotive interactions of help-giving, help-seeking and encouraging each other. Positive interdependence within a learning group can be established in different ways. The most prominent approaches have been positive outcome interdependence (i.e, setting group goals that learners must jointly achieve) and positive resource interdependence (i.e., giving learners independently access to different resources). Positive outcome interdependence alone already cause higher learning achievements compared to no outcome interdependence, but researchers showed that when outcome and resource interdependence are combined, groups in this condition outperform groups with just one type of interdependence (Johnson et al., 1989; Ortiz et al., 1996). Moreover, resource interdependence is associated with complementary viewpoints which produced more positive interactions among students (Buchs et al., 2004). However, the effects of positive outcome and resource interdependence are ambiguous. Buchs et al. (2004) emphasized that learners with positive resource interdependence only showed higher learning gains when information exchange was of high quality, but even then learning achievements did not always occur (Golub & Buchs, 2014).

As outlined above, searching and retrieving information can be considered a highly complex task. I argue that, together with positive outcome interdependence, positive resource interdependence will cause learners to generally engage more in promotive interactions, which will be particularly afforded by the ill-structured IPS task. Therefore, letting learners gain complementary knowledge prior to the search process should have a positive impact on groups' search activities. Similar to the approach of Dinet and Vivian (2012) and Lazonder (2005), the focus lies on the efficacy to find information. It is assumed that due to outcome and resource interdependence learners will discuss more relevant search terms and better coordinate their search activities, which results in a wider range of search terms on the group level. Consequently, a wider range of search terms will cause learners to scan a wider range of result items on group level. Due to a better information exchange and coordination, learners with resource interdependence will be better informed about the different aspects of the to-be-explored research topic, will know more technical terms during their search activities, and have a higher awareness for connected research fields. Already during scanning of result pages, learners can therefore better discriminate important from unimportant result items, and are more selective in regard to requests of details for specific result items. In turn, learners will also need less time to find relevant information within the details. In the present scenario, relevant information refers to compiled search items such as keywords and authors. Thus, positive resource interdependence and promotive interactions will increase a) the amount of different search terms entered, and b) the amount of different result items scanned, while it will decrease c) the amount of result items read as well as d) the total reading time within the group to complete the task (Hypothesis 1).

Moreover, promotive interactions and the exchange of information should stimulate cognitive processes. In their cognitive/elaboration framework, Dansereau and colleagues propose that exchanging information, as well as explaining or discussing task related concepts, activates cognitive structures, leads to reorganizations of cognitive structures, and helps to add new information into existing cognitive structures (Dansereau, 1988; O'Donnell & O'Kelly, 1994). As in individual elaboration, I predict that collaborative elaborations as part of promotive interactions will stimulate extra processing of relevant cognitive structures, hence of relevant keyword and author information (see the notion of elaboration in Reder, 1980). As learners under resource interdependence should engage more in discussions, and should exhibit more extra processing of the material, this in turn should lead to better recall rates for keyword and author information after the task (Hypothesis 2).

Even though positive resource interdependence should lead to more promotive interactions, and thus to better search performances and higher recall abilities, little is known about how learners reflect and perceive the learning process, or in particular, the learning partner (Bertucci, Johnson, Johnson, & Conte, 2012). Therefore, the concept of *self-other-agreement* was adapted from social perception theory to examine how learners perceive the learning partner in regard to their recall abilities after the collaboration.

Self-other-agreement

Research in social perception investigated how individuals perceive characteristics of other persons. Different methods have been applied to describe perception quality (Biesanz et al., 2007; Overall, Fletcher, & Kenny, 2012; Vazire, 2010; Watson et al., 2000). Kenny and West (2010) provided a distinct framework, in which the possible methods and procedures were categorized. Following their terminology, I am interested in a learner's (observer) perception of her partner (target), and adapt the two following measures: *self-other-agreement* between an observer's target perception and how the target perceives him-/herself; and *assumed similarity* between an observer's self-perception and his or her perception of the target person.

Even though self-other agreement is quite accurate for surface characteristics of a person (Fiske, 1993), it needs to be improved when people are prompted to evaluate hidden, underlying characteristics of a target person (John & Robins, 1993; Vazire, 2010; Watson et al., 2000). For example, it has been shown that people falsely believe that others hold the same knowledge as themselves, especially when they have very little information about the target person (Birch & Bloom, 2004; Nickerson, 1999). This means that generally people tend to assume a high similarity between themselves and others. This assumed similarity gets questioned the more information people receive from the target person. Letzring et al. (2006) demonstrated that both quantity and quality of exchanged information within naturally interacting dyads have an influence on self-other-agreement. Moreover, when pairs talk and exchange information, their perceived knowledge of each other increases rapidly (Reis et al., 2011). During the course of interaction, even randomly grouped pairs get familiar with each other, which causes an emotional closeness to the partner and a higher willingness to help (Biesanz et al., 2007; Korchmaros & Kenny, 2006; Watson et al., 2000).

According to the literature, increased information exchange is beneficial for a better understanding of a partner's knowledge. At the same time, people should be able to better predict their own recall rates, as the collaborative elaboration process makes own knowledge, as well as the differences between own and partner knowledge, more salient. Therefore, learners under resource interdependence are assumed to predict own recall rates more accurately than learners under no resource interdependence, which should be mediated by the amount of learners' promotive interactions (Hypothesis 3a). Similarly, learners under resource interdependence will predict partner recall rates more accurately than learners under no resource interdependence (Hypothesis 3b). This time, accuracy should be mediated by the amount of a partner's promotive interactions while being controlled for assumed similarity.

The present study

To examine the hypotheses stated above, dyads were requested to solve an authentic IPS task. For this, participants were required to take the role of a scientific collaborator, and were asked to use a computational search environment in order to explore a scientific research field together with their partner. The starting point for the search task was a scientific article about self-directed learning and autonomous machine learning, and was therefore set between psychology / education and computer science/ cognitive science. The task was regarded as accomplished once the dyad members achieved to collaboratively find 18 keywords or authors (or a mix of them), which show a relationship to the initial article with regard to its content, in order to better contextualize the initial article. Provided with the abstract of the initial article, participants were expected to read the text, mark keywords within the text, use these keywords as search terms throughout the search session and scan the result pages. Once they encountered an interesting result item by its title or author, they could open, read and extract new keywords/authors from the item's abstract. Hence, dyads were compared, who received either complementary or identical information as supplementary material with regard to the initial article. Receiving complementary material, dyads were assumed to individually gain complementary knowledge within the research field and to engage in more promotive interactions.

4.2 Method

Participants and Design

One hundred thirty-four students from a university in Germany participated in the study, receiving either 12 Euros or course credit as remuneration for the 90-min-experimental-session. In order to implement the complementary/identical viewpoints, only students from psychology, education, computer science and cognitive science in the third semester or higher were recruited. Comparing the curricula of all four academic subjects, students from psychology and education were treated as knowledgeable with regard to self-directed learning, while students from computer science and cognitive science were treated as knowledgeable with regard to machine learning. During recruitment, participants had to indicate their subject of study and were then randomly paired with either another participant having identical background knowledge (id-kl), or paired with a participant holding complementary background knowledge (com-kl). Then, participants received the supplementary material according to their background knowledge. This grouping resulted in 46 dyads with id-kl, 21.9 (SD = 2.48) years old on average, and 19 dyads with com-kl, 20.92 (SD = 2.31) years old on average. From the 46 dyads with id-kl 14 were gender-mixed dyads while 32 were same-gender dyads. In the other condition, eight mixed-gender and 11 same-gender dyads were encountered.

Material

Initial article. Through a comprehensive research in Google Scholar and Web of Science, articles were collected that dealt in equal parts with computational and psychological/ pedagogical issues. Eight articles were selected which fulfilled this requirement. In a preliminary study, 20 participants, 10 from the field of computer sciences and 10 from the field of psychology, rated the final article "Self-Directed Learning: A Cognitive and Computational Perspective" (Gureckis & Markant, 2012) as the best article according to overall relevance with regard to participants' own discipline (M = 4.80; SD = 1.01; 1 = *not relevant at all*, 6 = *very relevant*), according to legibility (M = 5.20; SD = 1.15; 1 = *not legible at all*, 6 = *very legible*), and according to understandability (M = 5.30; SD = 0.87; 1 = *not understandable at all*, 6 = *very understandable*). The article contrasts the notion of "self-directed" between human learning behavior and machine learning behavior (for a closer look, see Appendix B – Initial article).

Identical/ complementary material. Prior to the collaborative IPS phase, learners received a short encyclopedic text about either self-directed learning or machine learning. The texts were meant as baseline knowledge for all students and ideally stimulated participants' prior knowledge about the topic. Therefore, students from psychology/educaton always received a text about self-directed learning, while students form computer/cognitive sciences always received a text about machine learning. For the self-directed learning text, I extracted a passage from *The International Encyclopedia of Education* (1994), while the machine-learning text was composed by two English Wikipedia-articles ("Active learning", 2014; "Machine learning", 2014). Both texts contained 96-100 words and were provided in the original English version (see Appendix C - Encyclopedic texts about self-directed learning and machine learning).

Scientific Database. In order to explore the topic of the initial article, participants need to be provided with a scientific database similar to Google Scholar or Web of Science, in which they could discover new keywords or authors. As many books do not provide abstracts or a short introduction, the type of media was reduced to peer-reviewed journal articles. To further reduce the pool of potential search results, the initial article's cited references were scanned and journal names extracted, which had either a computational or psychological/ pedagogical scope. Then I completed the list to the extent that both topics were represented by 30 journals (see Appendix D - List of 60 scientific journal names for psychology/ learning sciences and computer/ cognitive sciences). In order to guarantee a stable database with all needed information (i.e. title, author, publication year, journal and abstract), newer articles from 2014 and 2015 were omitted, as information was sometimes missing. The set of 60 journals, with peer-reviewed articles until 2013, was the final database.

Devices. From a technical point of view, researchers from the information sciences have been exploring different devices and types of user-system interactions for collaborative information searching, which brought up solutions to use multitouch tabletops as the technical environment (Foley & Smeaton, 2009; Morris, Lombardo, & Wigdor, 2010; Morris, Paepcke, & Winograd, 2006; Pickens, Golovchinsky, Shah, Qvarfordt, & Back, 2008). As my goal was to provide participants with a device which allowed simultaneous interactions with the search system, while preserving intragroup face-to-face communication, multitouch tabletops were adopted as the technical environment for this quasi-experiment. In contrast to the literature mentioned above, no further tools were implemented to better identify the impact of promotive interactions on learners' search activities, which were logged by the system. In addition to the multitouch tabletop, two laptops were used to ask participants separately for their demographic data and to test their recall abilities at the end of the quasi-experiment.

Interface. I designed a search interface (Figure 4), which was similar to commonly known search engines on the Internet. Thus, participants' cognitive load was reduced by activating their preexisting mental representations of such search engines. The design provided for each participant one search window, which was built of two tabs. One tab hosted the search functionality, while the second one listed all bookmarked search results. The "search tab" consisted of a text entry field in the head of the tab to enter search terms, and a frame below where result items were presented list-wise.

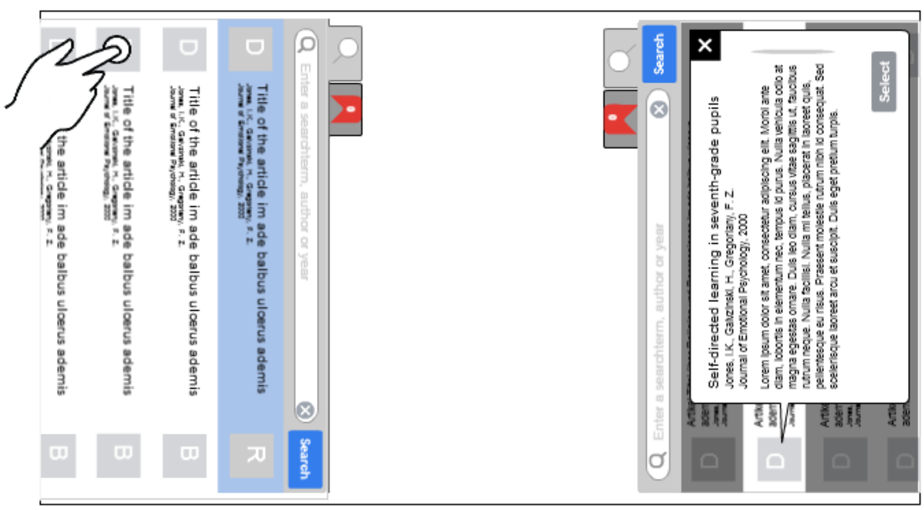

Figure 4. Multitouch tabletop search interface for two persons
Interface showing on the left side the search tab with a result list, of which the first item was bookmarked and where details are requested for the third item. On the right side, details of the second item were requested and an overlay window shows the information including title, author, year, journal name and abstract.

To type search terms into the text entry field, participants were provided with a virtual keyboard using the German qwertz-design. Lists contained up to ten result items at a time so that additional result items were split into several pages, to which users could jump by using a pager (i.e. backward and forward arrows with a page counter). Result items were represented by title, author, year, journal name and an initially hidden abstract. The ordering of result items in the list was based on a predefined relevance ranking, which took the amount of search terms' appearances in title, author, year, journal name and abstract into account. To open the abstract of a result item or to bookmark interesting result items, two buttons were implemented for every result item in the list. The "details"-button opened an overlay window, showing the complete information of the result item including its abstract, while the "bookmark"-button colored the result item in the list and duplicated the item automatically into the bookmark-tab list. The "bookmark tab" contained a frame listing all bookmarked articles ordered by the time they were marked. There, bookmarked items could be re-examined by opening the details or by removing the item in case it lost its relevance during the search process.

Measures

Condition. As mentioned earlier, I experimentally manipulated whether dyads have identical knowledge or complementary knowledge prior to the collaborative IPS phase. This variable was always treated as a factor with two levels and coded -1 (id-kl) and +1 (com-kl) for further analysis.

Amount of search terms / Amount of scanned items / Amount of read items / Total amount of reading time*.* To analyze whether dyads with complementary knowledge *search* with a higher diversity of search terms, all search terms were pooled on group level which were entered by at least one of the two participants. Search terms were treated as different when participants added or removed words of an initially entered search term to either increase or decrease the amount of result items. For instance, 'learning' was seen as a solitary derivative from 'self-directed learning', as well as 'self-directed learning children'. However, misspellings, reordering of words within a search term or translations of search terms from English to German were not treated as different in this regard, as they represent no elaborations whatsoever. Finally, search terms which occurred several times after being pooled, were counted only once. Next, I looked on the amount of *scanned* result items. Result items were treated as scanned when they were listed in a result page, which was opened by either one participant or the other. Like with search terms, result items were pooled on group level and cleared for redundancies. Similarly, all result items were collected, whose abstracts were opened and *read* for at least 5s by either one or the other participant. These items were also cleared for redundancies after being pooled. Finally, I summed up the *total amount of reading time* in minutes of both dyad members for requested detail information, which was opened longer than 5s.

Self-reported discussion (SRD). After the collaborative search process, participants were asked separately to judge their collaborative elaboration of keywords and authors: "Would you agree with the following statement? For me it was more important to discuss every keyword/author with my partner than just quickly search and select keywords or authors." (1 = *I totally disagree*, 7 = *I totally agree*). Responses were used as a discrete variable for further analysis.

Prediction of own recall (PoOR) / *prediction of partner recall* (PoPR). After participants selected their final list of 18 keywords/authors, each participant's prediction of recall (PoOR) of every single keyword or author was assessed, as well as their estimates of partner's recall abilities for each item (PoPR). Therefore, I used a question from the metacognition literature (e.g. Dunlosky & Lipko, 2007; Thiede et al., 2003), displayed it below a selected keyword/author ("How well will you/your partner be able to recall this selected keyword/author in a test in about 10 minutes?"), and asked participants to answer the question by entering an integer between 0 and 100 in a text field (0 = *very badly*, 100 = *very well*). Afterwards, the eighteen self-predictions and partner-predictions were averaged.

Recall test (RT). Finally, recall test scores were calculated by noting the percentage of correct recalled keywords/authors in the final test from 0 to 1. Keywords were correctly recalled also when errors of spelling or upper/lower cases occurred. Also when participants did not recall the forename of an author, the right last name was counted as a correct recall. Then, performance scores in percentiles were stored as continuous variables for both dyad members.

Procedure

Arriving at the lab, the experimenter welcomed both participants and let them introduce themselves including their subject of study so that each learner's discipline became salient to the other person. Then, participants were seated in different cubicles with a laptop and were asked for their demographic data. After finishing the questionnaire, participants were handed a clipboard with their task, the initial abstract and the short, encyclopedic text on separate sheets of paper. The task encompassed the scenario of scientific collaborators and the goal to search and retrieve 18 keywords or authors from other related articles, which show a relevance to the initial article. Additionally, participants received information on a potential search strategy (compare "process worksheets"; Hummel, Paas, & Koper, 2004). The overall time limit was set to 50 minutes. Participants read the task, as well as the priming text, and were allowed to ask questions concerning details of the task. Comprehension questions about the original article or the encyclopedic texts were not answered. Afterwards, the experimenter invited both participants to the multitouch tabletop. Participants were instructed to take their clipboards with them so that they could always have a second look into the task and use the back of a sheet to write down selected keywords/authors. At the tabletop, the experimenter introduced the search interface to the participants and demonstrated all functionalities. Using two unrelated search terms, participants could then test the functionalities on their own for a maximum of five minutes to get familiar with the interface. Following this test phase, the participants began to solve the task, while the experimenter left the room leaving an English-German dictionary to the participants in case of language difficulties. After 35min., the experimenter first re-entered the lab to remind participants of the last 15min. All dyads achieved to select 18 keywords/authors within the given time frame, although the majority needed the complete 50min to succeed. Together, the learners entered all keywords/authors into a form back at their laptops. After entering the keywords/authors, participants were separated again and seated into their cubicles to complete the last part of the experiment. First, they were asked to provide for each keyword/author a prediction of their own and their partner's recall ability. Afterwards, they read an unrelated text to clear working memory and reported their discussion attitude during the IPS task. Finally, both participants tried to freely recall as much keywords/authors as possible at their own pace. Once they submitted their recalled list, the experiment ended and participants were remunerated.

Analysis

For hypothesis 3 (prediction of own and partner's recall performance), I used the truth-and-bias model of West and Kenny (2011) as the general paradigm. The paradigm uses the estimation of recall abilities as the outcome variable, while the respective performance value is treated as the predictor, called truth variable. This way the model allows implementing bias variables as supplementary predictors, which can take the role of a mediator or control variable. For the first part of the hypothesis, for which the association between a learner's actual recall performance (i.e. truth variable) and his/her prediction of recall was analyzed, I used a learner's self-reported discussion attitude as a potential mediator (Figure 5). In the second part, the association between the partner's actual recall performance (i.e. truth variable) and the prediction of the partner's recall performance is analyzed, while the partner's self-reported discussion attitude was used as a potential mediator, and the prediction of the learner's own recall performance was used as a control variable for assumed similarity (Figure 6).

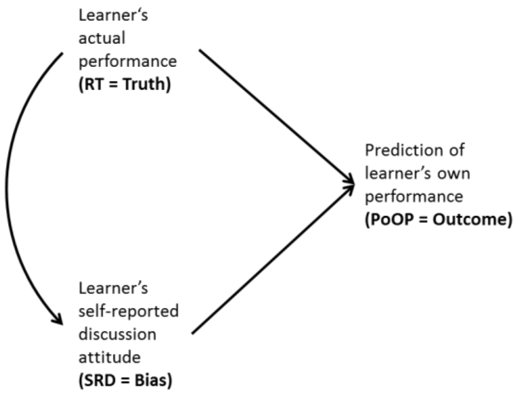

Figure 6: Model of absolute self-monitoring accuracy in Study 3.
Truth-bias-model estimating the accuracy of learner's own recall prediction, potentially mediated by learner's self-reported discussion attitude during the cooperative IPS.

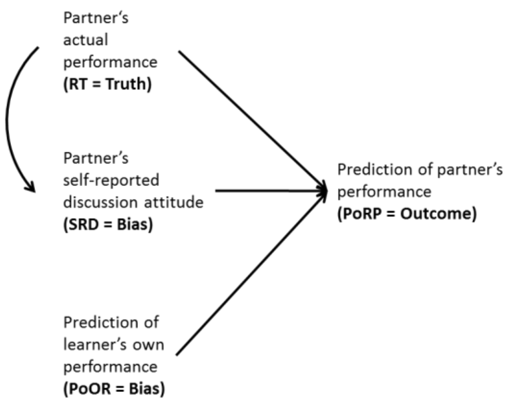

Figure 5: Model of absolute partner-monitoring accuracy in Study 3.
Truth-bias-model estimating the accuracy of learner's prediction of partner performance, potentially mediated by partner's self-reported discussion attitude and controlled by learner's prediction of own performance.

However, dyad members are not independent of each other, so that the influence of one member on the other has to be taken into consideration. Therefore, all analyses which were not on group level, such as the truth-bias-model, were computed using the actor-partner-interdependence-model (APIM: Cook & Kenny, 2005; Kenny, Kashy, & Cook, 2006). The APIM handles the specific situation of interdependent dyads and always takes the partner into account (i.e., partner effect) when estimating a learner's influence on the outcome variable (i.e., actor effect). Therefore, estimates of a learner's outcome variable can always be separated into actor and partner effects. For example, the association between a partner's actual recall performance and the learner's prediction of partner performance (outcome variable) is a partner effect, while the association between the learner's actual performance and her prediction of performance (outcome variable) is an actor effect. APIM uses general least square analysis with correlated errors and restricted maximum likelihood estimations to calculate this multi-level approach (R-script "APIMM.R" provided by David Kenny and Thomas Ledermann[3]).

4.3 Results

In H1 it was predicted that dyads with complementary knowledge show more efficient search actions than dyads with identical knowledge. More specifically, I assumed that due to promotive interactions, dyads with complementary knowledge will show a higher amount of search terms, which should subsequently lead to a higher amount of different result items scanned. Moreover, promotive interactions enable learners to be more informed and selective for interesting and relevant items so that the amount of items read as well as the total reading time should be decreased.

However, the statistics using independent t-tests give no evidence that dyads with com-kl are different from dyads with id-kl. Dyads with id-kl used a similar amount of search terms (M = 14.52, SD = 7.37) as dyads with com-kl (M = 14.89, SD = 6.71); t(62) = 0.18, p = .856, r = 0.023. Dyads with id-kl also scanned a similar amount of result items (M = 186.35, SD = 87.27) and therefore paged through a similar amount of result lists as dyads with com-kl (M = 224.89, SD = 100.63); t(62) = 1.51, p = .137,

[3] Retrieved the 16th of June 2015; By date of submission, an updated web-based app permits the computation on https://davidakenny.shinyapps.io/APIM_MM/ , which reveals the same results like the R-Script

r = 0.188. Further, dyads with id-kl read a similar amount of result items (*M* = 19.93, *SD* = 7.53) as dyads with com-kl (*M* = 19.56, *SD* = 7.36, t(62) = -0.18, *p* = .857, r = 0.023), and also the reading time did not differ significantly between dyads with id-kl (*M* = 36.55, *SD* = 11.10) and dyads with com-kl (*M* = 32.02, *SD* = 11.09); t(63) = -1.49, *p* = .142, r = 0.184. Even though the hypotheses need to be rejected, it can be stated that effect sizes for reading time point out a small effect in the hypothesized direction.

Concerning each learner's recall rates (RT) of compiled keywords and authors, H2 predicted that learners in dyads with complementary knowledge will recall more items than learners in dyads with identical knowledge. Moreover, I predicted that learners will be positively influenced by their promotive interactions within the dyad so that a positive relationship between self-reported discussion attitude (SRD) and recall performance should occur only for learners in dyads with com-kl. Therefore, the actor effect of SRD on RT was of particular interest and was calculated by an actor-partner-interdependence-model (APIM), which is presented in Table 11.

Table 11. Actor effects of one APIM testing the moderation of condition within the association between self-reported discussion attitude (SRD) and recall test performance (RT; H2). Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and effect size (r).

	Δ R^2	B	SE	*β*	p	r
RT						
Intercept		0.50*	0.02		<.001	
Condition		-0.01	0.02	-0.05	.643	-.048
SRD (actor effect)		-0.01	0.01	-0.06	.553	-.054
SRD x Cond.		-0.03*	0.01	-0.21	.036	-.190
Model	.039					

The intercept indicates that learners achieved to recall nine keywords/authors on average (50 %), while the predictor 'condition' indicates no difference between resource-independent or resource-interdependent groups. Looking at the relationship between SRD and RT, it can be seen that the relationship negatively interacts with the condition. Therefore, I ran two additional APIMs to analyze the relationship for each condition separately using a dummy coding (i.e. 0|+1 to estimate effects of dyads with id-kl and +1|0 to estimate effects for dyads with com-kl). Results revealed that recall of dyads with id-kl were positively, though not significantly, associated with their discussion attitudes (B = 0.02, SE = 0.01, β = 0.15, p = .154, r = .128), whereas recall of dyads with com-kl were negatively, but not significantly, associated with their discussion attitudes (B = -0.03, SE = 0.02, β = -0.26, p = .112, r = -.146). Thus, against the stated hypothesis the two conditions show similar recall rates, while post-hoc tests showed that dyads with id-kl, rather than dyads with com-kl, benefit from discussions.

H3 looks at learners' prediction accuracy to generally estimate own as well as partner recall rate after the problem-solving phase. It was argued that members of dyads with complementary knowledge should benefit from increased promotive interactions, exchange more information and thus have a better awareness about own and partner knowledge characteristics. Hence, they should also predict own and partner recall rates more accurately than members of dyads with identical knowledge. As outlined in the methods section, the truth-and-bias model of West and Kenny (2011) was used as the general approach to analyze prediction accuracy. Following the authors' recommendations, all variables were centered so that the intercept of the two models directly indicates whether participants over- or underestimate themselves or the partner. For this, in both models Figure 5 and Figure 6 prediction and performance variables were centered by the respective grand mean of the truth variable, while self-reported discussion variables were centered by their own grand mean. Positive intercepts point therefore to an overestimation, whereas negative intercepts indicate an underestimation when all predictors are averaged. To test H3a by calculating the associations in Figure 5, two APIMs were used. The first APIM uses learners' own recall test performance (RT) and their self-reported discussion attitude (SRD), as well as the variables' interaction terms with condition (id-kl vs. com-kl) as predictors, while prediction of own recall (PoOR) was the outcome variable. In the second APIM, RT was used as the predictor and SRD as the outcome variable. Results are presented in Table 12.

Table 12. Actor effects of the first APIM tested the moderation of condition within the associations between recall test performance (RT) and prediction of own recall (PoOR; outcome variable), and between self-reported discussion attitude (SRD) and PoOR. The second APIM tested the moderation of condition within the associations between RT and SRD (outcome variable).
Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and effect size (r).

	Δ R^2	B	SE	*β*	p	r
PoOR						
Intercept		0.03*	0.01		.026	
Condition		-0.03	0.01	-0.20	.025	-.214
RT (actor effect)		0.22*	0.08	0.28	.010	.231
RT x Cond.		-0.07	0.08	-0.09	.412	-.074
SRD (actor effect)		0.00	0.01	0.02	.863	.016
SRD x Cond.		0.01	0.01	0.08	.409	.076
Model	.116					
SRD						
Intercept		-0.06	0.15		.690	
Condition		-0.11	0.15	-0.06	.461	-.066
RT (actor effect)		-1.24	0.86	-0.15	.153	-.128
RT x Cond.		-1.76*	0.86	-0.22	.043	-.181
Model	.040					

The intercept of the first APIM shows that participants generally overestimate themselves by around three percentile points. However, overestimation decreases significantly for dyads with com-kl (indicated by the 'condition' estimates). Therefore, H3a can partly be confirmed. To see if promotive interactions caused dyads with com-kl to be more accurate in predicting own recall performance, the remaining estimates are examined. The moderation effect on the relationship of learners' recall rate with the outcome variable shows no difference. But as the general association became significant, it can be stated that all learners were affected by their own knowledge when predicting their recall performance. Examining the interaction term for learners' self-reported discussion attitude (predictor), no difference is found between dyads with id-kl and com-kl, and also, generally spoken, learners did not align higher recall abilities with higher discussion attitudes and vice versa. Thus, promotive interactions seem to have no influence on self-perception accuracy so that learners' higher accuracy rates might be influenced by other factors. As SRD is not a significant predictor for learners' prediction of own recall, I exclude further reporting of the path between RT and SRD.

To test H3b, calculating the associations in Figure 6, two APIMs were computed. The first APIM used partners' actual recall rate (RT), partners' self-reported discussion attitude (SRD) and learners' prediction of own recall (PoOR; to control for assumed similarity), as well as their interaction terms with the condition, as predictors. Learners' prediction of partner recall (PoPR) was taken as the outcome variable. In the second APIM, the partner's actual recall rate was the predictor and the partner's self-reported discussion attitude the outcome variable. The results are presented in Table 13.

Table 13. Actor effects of the first APIM tested the moderation of condition within the associations between partner's recall test performance (RT) and prediction of partner's recall (PoPR; outcome variable), between partner's self-reported discussion attitude (SRD) and PoPR, and between learner's prediction of own recall (PoOR) and PoPR. The second APIM tested the moderation of condition within the associations between partner's RT and partner's SRD (outcome variable).
Indicated are adjusted R^2, estimates (B), standard error (SE), standardized regression coefficients (β), p-values and effect size (r).

	Δ R^2	**B**	**SE**	***Β***	**p**	**r**
PoPR						
Intercept		0.05*	0.01		<.001	
Condition		-0.01	0.01	-0.03	.472	-.061
RT (partner effect)		0.06	0.04	0.08	.199	.121
RT x Cond.		-0.07	0.04	-0.10	.110	-.150
SRD (partner effect)		-0.01	0.00	-0.10	.060	-.173
SRD x Cond.		-0.01	0.00	-0.08	.139	-.137
PoOR (actor effect)		0.73*	0.05	0.81	<.001	.792
PoOR x Cond.		-0.06	0.05	-0.07	.230	-.113
Model	.746					
Partner's SRD						
Intercept		-0.06	0.15		.690	
Condition		-0.11	0.15	-0.06	.461	-.066
RT (partner effect)		-1.24	0.86	-0.15	.153	-.128
RT x Cond.		-1.76*	0.86	-0.19	.043	-.181
Model	.040					

The intercept turned out to be significantly positive showing that generally participants overestimate their partner's performance by around five percentile points, while no difference occurs between the two conditions. Thus, against the hypothesis, dyads with com-kl are not more accurate in estimating their partner's recall rates than dyads with id-kl. Furthermore, partners' SRD did not show any association with predicting partner recall rate, neither in dyads with com-kl nor in dyads with id-kl. However, learners were not directly affected by their partner's recall abilities, as the non-significant estimates for the predictor RT and the moderation effect of RT with the condition demonstrate. Rather, learners were heavily influenced by their PoOR and this was irrespective of whether learners were grouped with partners of com-kl or id-kl. In other words, learners generally assumed a high similarity between their own and their partner's recall abilities independently of the groups' knowledge constellation prior to the collaborative session. As partners' recall rate and partners' discussion attitude show no influence on the prediction of partner performance, I exclude further reporting of the path between partner recall rate and partner discussion attitude. Taken together, H3b has to be rejected as dyads with com-kl were not more accurate in predicting partners' recall rates than dyads with id-kl, and they were also not influenced by partners' promotive interactions.

4.4 Discussion

The goal of this chapter was to transfer and replicate positive effects of resource interdependence and promotive interactions from the cooperative learning literature to the cooperative information problem solving process. For this, dyads were asked to search and compile keywords/authors related to an interdisciplinary research field, while the material that dyads received prior to the cooperative phase (identical vs. complementary) was manipulated. First, dyadic search efficiency was examined. Second, dyads' recall ability of compiled keywords/authors was tested. And finally, dyad members' capability of accurately predicting own and partner recall rates was investigated. In asking for their predictions, the focus lied on dyad members' representations of own and partner knowledge, and on the possible influence that both representations have on each other.

Results revealed that the replication of positive effects of complementary prior knowledge was not successful. First, dyads with complementary material (i.e., under positive resource interdependence) were not more efficient in their search activities than dyads with identical material. Even though total time on task could not be examined due to the fixed time limit, dyads with identical material used a similar amount of search terms, scanned a similar amount of result pages and read a similar amount of requested abstracts. However, the total reading time tends to show the expected pattern of efficiency with small effect sizes, which indicates that dyads with complementary material might be more focused and might be better at filtering relevant information from the content compared to dyads with identical material.

Nonetheless, clear evidence with large effect sizes for a higher efficiency in dyads with complementary material has not been found.

Second, dyads with complementary material showed no improved recall rates compared to dyads with identical material, and therefore this aspect of group performance did not bring up any evidence that dyads with complementary material activated more cognitive structures during the collaboration than dyads with identical material. As both conditions yielded a similar average recall (9 items out of 18), it implicates that other factors influenced this recall rate, which were equal for both conditions. These could be factors apart from collaborative problem solving process, such as the type of final presentation (collaboratively entering the compiled items), or the time delay between entering the keywords and the free recall. Therefore, excluding a collaborative entering of compiled items might better elicit differences in dyads' actual recall abilities.

With regard to learners' self- and other-perceptions after the IPS task, members of dyads with complementary material showed higher prediction accuracy of own recall rates than members of dyads with identical material. While learners with identical material overestimated themselves by three percentile points, learners with complementary material were perfectly accurate. However, the effect size was rather small and the overestimation of dyads with identical material could be also just an error of measurement. While predictions were influenced by learners' actual recall rate, indicating that learners had some insight into their own knowledge, predictions were not influenced by learners' amount of engagement in discussions. With regard to partner prediction accuracy, all participants overestimated their partner's recall performance by approximately five percentile points, which can be considered a good insight into partners' mental states. But as predictions of partner performance was neither influenced by partners' actual recall rate, nor by partners' engagement in discussions, it seems that learners transferred their predictions of own recall to their partners, which is suggested by learners' heightened assumed similarity (Birch & Bloom, 2004).

In contrast to the interpretation that learners have a good insight into their own and their partner's recall abilities, a second hypothesis might also explain learners' high prediction accuracies, which is provided by Kruger and Dunning (1999). The authors conducted a series of studies in regard to over- and underestimation. They found that people tend to use a performance between 40-60 % as a desirable anchor to judge themselves towards this anchor. In the present study, this anchor yielded some precision given that all participants achieved around 50 percentile points in the recall test. Therefore, learners might have predicted own performance rates according to this desirable anchor and also transferred this anchor to their partners. However, this usually implicates that learners do not take much further information into account when predicting their own knowledge state, but base their predictions solely on general beliefs with such recall tests. This might also explain why discussions within the group, whether they occurred or not, did not have any influence on learners' predictions.

Different reasons might have contributed to the result that positive resource interdependence had no impact on dyads' search processes or cognitive performances. First, it can be questioned if the task contained the desired complexity or difficulty to strengthen learners' motivation to interact during the IPS task. In fact, students are highly familiar with search engines like Google and are used to move through hypertext environments so that this non-linear environment did not necessarily increase learners' need for help or willingness to exchange information with their partners. Second, in order to compile the 18 items it was not mandatory to share every finding and discuss every compiled item so that dyads could have also simply divided the labor and continued to search and compile on their own. However, being tied together concerning the overall goal, as well as being confronted with an interdisciplinary subject and with a collaborator with complementary knowledge, was expected to make promotive interactions almost inevitable. Saying that, the task did not explicitly require participants to actively integrate prior knowledge or compiled items in order to uncover for example a hidden best selection (compare to Stasser & Titus, 2003), or the solution of a problem with one exact answer (compare to Aronson, 1997). Third, I implemented a specific topic and recruited a sample that might not be highly generalizable to other topics and samples. As multidisciplinarity is nonetheless a very important topic in higher education, as well as in the professional context, further interdisciplinary topics have to be tested to validate or refute these findings. Finally, it has been shown that groups sometimes need additional triggers to actually engage in promotive interactions such as discussion rules in constructive controversy (Golub & Buchs, 2014). Therefore, more effort has to be put into identifying the exact circumstances under which dyads need to exchange and integrate their complementary knowledge in order to increase search efficiency and quality.

Even though the manipulation of positive resource interdependence did not affect dyadic search or cognitive behavior, the necessity to investigate more deeply the information problem solving process of groups needs to be emphasized. Students need to complete more and more group projects, whether in college or in higher education. Therefore, understanding the dynamics of cooperative learning groups with an informational need is a crucial aspect to develop effective instructional designs and technological support in order to increase better problem-solving performances and higher individual knowledge gains. Moreover, a method was adapted from social perception, which gave more insight into the cognitive processes of each dyad member. It revealed for example how strongly people assume a similarity between themselves and their partners, especially when interactions between group members have no great value.

Therefore, upcoming studies need to investigate more deeply how positive resource interdependence might increase promotive interactions during the IPS process. For this, the ill-structured task has to be revised so that learners are more likely to exchange and discuss prior as well as newly retrieved information. At the same time, the ill-structured design needs to be compared with more well-structured designs including close-ended search environments and a clear set of to-be-retrieved items. Additionally, other (interdisciplinary) topics have to be tested to achieve more generalizability and to exclude effects of participants' personal (dis)interest in the topic. Moreover, the aspect of individual accountability within cooperative learning should be considered. In the present study, all participants were equally motivated to solve the task (by the provided remuneration), but specific motivational incentives, such as epistemic motivation (Scholten, van Knippenberg, Nijstad, & De Dreu, 2007), could increase learners' individual accountability and lead to more promotive interactions (Johnson et al., 2007). Nonetheless, to correctly analyze the information exchange within the dyad, the actual utterances need to be captured using coding methods to identify which types of utterances had an influence on the problem-solving process or on the cognitive structures (Baker et al., 2007; Dehler et al., 2011). Finally, from a cognitive perspective it has to be examined when learners start differentiating between own knowledge and partner knowledge. In the present scenario, I asked learners to predict their partner's knowledge with regard to their joint final product. But how would predictions look like if learners were asked to predict partners' prior knowledge and how good partners might answer to questions in regard to self-directed learning or respectively to machine learning? This might be especially important when learners have to rely on partners' prior knowledge, or even have to learn from them. Therefore, tasks should be taken into consideration which specifically encourage learners to gain expert knowledge independently, which has to be combined to solve the task, and then to analyze whether learners can monitor and control their partner's knowledge.

Chapter 5

General Discussion

Learning in groups, whether self-organized or carefully directed by the teacher, has proved its impact on individual learning over the last fifty years. However, cooperative learning is influenced by many individual and group level factors, and clear recommendations for instructional and technical support are difficult to draw as research results on e.g. group constellation, motivational factors or instructional rules are ambiguous. At the same time, more and more project-based techniques are used in school and higher education to strengthen learners' problem-solving, communication and cooperation skills. Scholars need to better understand how groups regulate their learning activities in order to provide them with instructional and technical support to maximize their learning achievements. As one of the first in this field, the aim of the present dissertation was to provide the bases to study the regulative mechanisms of cooperative dyadic learning from an individual metacognitive perspective. For this, I developed an extension of Nelson and Narens' (1994) object-metal-level model for dyadic interactions (MEDIA; Figure 1, page 21), which intertwines cognitive and social behavioral processes. Therefore, it can be assigned to the field of social metacognition (Jost et al., 1998). The model explicates many different assumptions, from which some have been tested in the empirical studies presented here. In the following section, I will first summarize the main findings study by study, and will then try to give an overview of confirmed and disconfirmed assumptions within MEDIA by separating absolute from relative measurements, especially in regard to monitoring and control accuracy.

5.1 Summary of the main findings

The first study presented in Chapter 2 put learners into a fictive cooperative scenario with a real individual learning phase and asked learners afterwards how much information they would provide and how much information they would expect from an ostensible learning partner during the course of a collaborative learning phase. To create (fictitious) homogeneous and heterogeneous dyad constellations, learners' assessments of knowledge and those of the ostensible partner were manipulated. First, the results provide evidence that learners indeed hold own and partners' individual discrepancies in mind, as the intentions to provide and the expectations to receive information depended both on own and on partners' evaluated knowledge discrepancy. The higher participants judged their own knowledge (i.e. low individual discrepancies), the more they were willing to give information, and the higher the ostensible learning partner judged her knowledge, the more learners expected information from their partner. Against my assumptions, learners' intentions and expectations were independent from intragroup differences. In other words, irrespective of how high ostensible partners judged their knowledge level, learners intended to give information when they judged themselves as high-ability learners, and irrespective of how high they estimated their own knowledge level, learners expected to receive information as long as partners (fictitiously) judged themselves as high-ability learners. However, in the more fine-grained measures of relative group control accuracy (see intercepts of Table 3 and Table 4) evaluations of intragroup differences have been found with small effect sizes, as learners tended to adjust their intentions and expectations of each learning item according to the extent of the respective intragroup difference. Second, no relationship between relative monitoring accuracy and relative control accuracy of providing/receiving explanation has been found, as it was proposed by MEDIA. That means, even though participants showed some relative control accuracy by taking intragroup differences into account, discriminating between well and poorly understood learning items had no influence on relative control accuracy.

The results concerning intentions and expectations of information exchange mirror findings from small group learning that high-ability learners in heterogeneous groups provide mostly unilateral explanations whereas in homogeneous groups they engage in mutual elaborations (Saleh et al., 2005; Webb, 1991). Consequently, low-ability learners mostly receive explanations in heterogeneous groups, while they rarely gain knowledge in homogeneous groups (Lou et al., 1996; Saleh et al., 2005). In this regard, the results suggest a cognitive rule of learners: as long as individual learning goals are not reached yet, learners will try to maximize their knowledge by engaging in mutual knowledge exchange independently of individual discrepancies and intragroup differences.

From an instructional perspective, the findings underscore that learners of mixed abilities should be grouped together, matching intentions and expectations of learners in an optimal way. In a similar scenario with an individual and a subsequent collaborative learning phase, Chapter 3 reported on an empirical study with real dyadic interaction. First, the study tried to replicate and differentiate findings from Chapter 2 concerning individual discrepancies. It has been found that high-ability learners provide more explanations than low-ability learners. In contrast, low-ability learners do not use many questions as means to elicit explanations. Second, the study could further replicate that learners take intragroup differences into account when it comes to relative group control accuracy, but only for high-ability learners with high monitoring accuracy (relative control accuracy for providing explanations in Table 6). Additionally, a relationship between relative self-monitoring accuracy and relative group control accuracy occurred, indicating that monitoring can indeed influence group control behavior on the item level. More specifically, when high-ability learners (not low-ability learners) are good in discriminating well and poorly understood learning items they can transfer this awareness to group interactions in a way that makes explanation-giving more efficient (i.e. providing more explanations to items with a positive knowledge difference and fewer explanations to items with a negative knowledge difference). For this to occur, evaluating intragroup differences was mandatory requirement.

Third, with the study I wanted to explore how control behavior influences self- and partner-monitoring, looking on absolute measures of monitoring accuracy. On the one hand, the results indicate that high-ability learners monitor themselves more accurately than low-ability learners, but that both types of learners are not affected by the amount of provided cues. On the other hand, both low- and high-ability learners monitor their partners equally well, and they are also influenced equally strongly by the amount of received cues and judgments of own comprehension. However, effect sizes point to the fact that learners tend to align their judgments of partner knowledge more with judgments of own knowledge than with the amount of received cues, exhibiting therefore a strong false consensus effect. Moreover, the amount of received cues was not mediating partner-monitoring accuracy so that it can be questioned whether received cues helped learners to judge their partners accurately or not.

Finally, as individual metacognition proposed that higher control accuracy will lead to higher learning achievements (Thiede et al., 2003), the study tested whether low-ability learners could benefit from high-ability learner’s efficiency to provide explanations (looking once again at relative measures of group control accuracy). Surprisingly, no association could be found in this regard. Rather, an association showed up between the amount of received explanations and low-ability learners’ knowledge gains.

Instead of investigating knowledge differences within a simple learning session, an authentic problem-solving task was presented to dyads in Study 3. Here, it was examined whether complementary prior knowledge within the dyad causes learners to engage in more interactions, whether this engagement is reflected by dyads' actual problem solving activities (as it is proposed in cooperative learning, Aronson, 1997; Johnson et al., 1989; Ortiz et al., 1996), and whether more interactions cause both learners to memorize and represent the problem solution more accurately. Learners needed to search and compile lists of topically relevant keywords and authors with regard to an interdisciplinary research field. It was expected that holding complementary knowledge with regard to the research field (in contrast to identical knowledge) should increase knowledge exchange and elaborations about potential keywords/authors. In turn, an increased amount of dyad interactions might also increase learners' awareness of own and partners' cognitive states. Similar to Study 2, the question was whether the amount of learners' interactions might affect self-monitoring accuracy of the memorized keywords and authors (i.e. solution), and whether the amount of partner elaborations might affect partner-monitoring accuracy of memorized keywords and authors. Analyzing absolute measures of accuracy, results revealed that learners in the complementary knowledge condition were slightly better in monitoring own recall abilities, while learners in both conditions slightly overestimated partner recall abilities. However, no influence of learners' elaborations was found on judging own recall abilities, and no influence of partner elaborations was found on judging partners' recall abilities. With regard to these cues, the results accompany findings from Study 2, which also showed that provided and received cues play no powerful role in self- and partner-monitoring accuracy. Moreover, judgments of learners' own recall abilities strongly influenced judgments of partner recall abilities, thus representing the same strong false consensus bias that was found in Study 2.

Together, several assumptions of MEDIA could be confirmed, while others need further examination. As noted, results can be interpreted in absolute or relative terms, e.g. when reporting the impact of provided explanations on learning gains (absolute measure) or when reporting the relative difference of well and poorly monitored learning items on the respective behavior of giving or eliciting explanations for these learning item (relative measures). Therefore, I first interpret the results in Figure 7 in absolute terms. Starting on the object-level, learners were able to represent self- and partner-representations, as in all studies participants needed to judge or react to these representations in different ways (Studies 1-3). Moreover, learners were specifically asked to separately assess own and partner level of understanding after the collaboration (Study 2; and in Study 3 item-by-item) on the meta-level. However, judging one's partner is strongly influenced by judging oneself and can therefore bias the correct evaluation of individual discrepancies (i.e. false consensus bias; Study 2; and also in Study 3 when measured on item level).

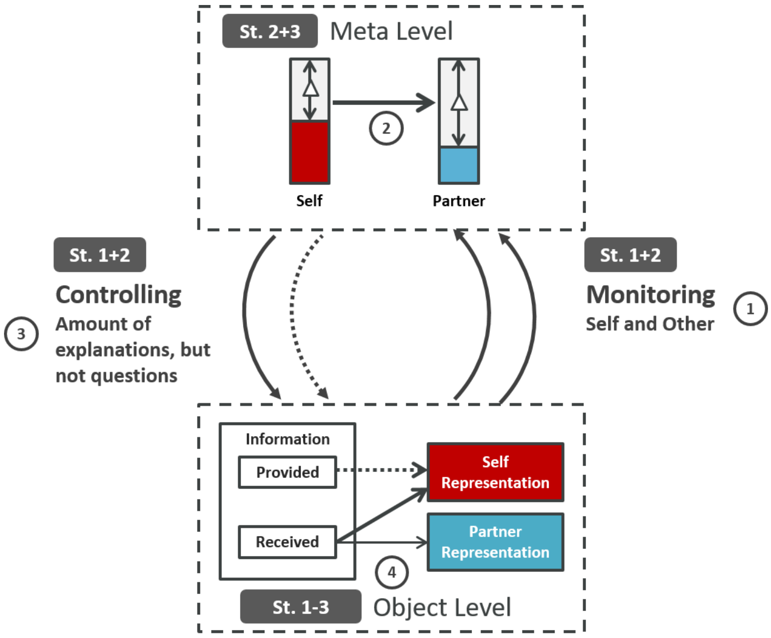

Figure 7. Results of the studies with regard to MEDIA focusing on absolute measurements
1. Learners could represent and monitor self- and partner-representations. 2. Judging and evaluating partner's knowledge (partner's individual discrepancy) was strongly influenced by learner's evaluation of own knowledge, 3. Studies in chapter 2 and 3 showed that high-ability learners used explanations as a mean to provide information, but that low-ability learners used rarely questions as a mean to elicit explanations, 4. Provided cues had no influence on the self-representation, while received cues had a small effect on learner's judgment of partner knowledge; the amount of received cues were correlated with knowledge gains

Study 1 showed that learners intend to provide explanations and expect to receive explanations, which has been confirmed during actual knowledge exchange (Study 2), but that they do not use questions to elicit them (Study 2). Coming back to the object-level, results concerning the influence of provided and received cues on self- and partner-monitoring accuracy are more ambiguous. While provided cues had no influence on self-monitoring accuracy (Study 2 and 4), the amount of received cues had a small effect on learners' judgments of partner knowledge, but did not mediate overall partner-monitoring accuracy (Study 2). Analyzing individual learning gains in Study 2, the study revealed that the total amount of received explanations had an influence on low-ability learners' knowledge gains. Contrary, no knowledge gains have been found for high-ability learners.

Second, Figure 8 summarizes the results in terms of relative measures. Beginning with the monitoring process, the empirical studies assessed only relative self-monitoring accuracy, leaving the ability to assess partner knowledge on item level as an open question.

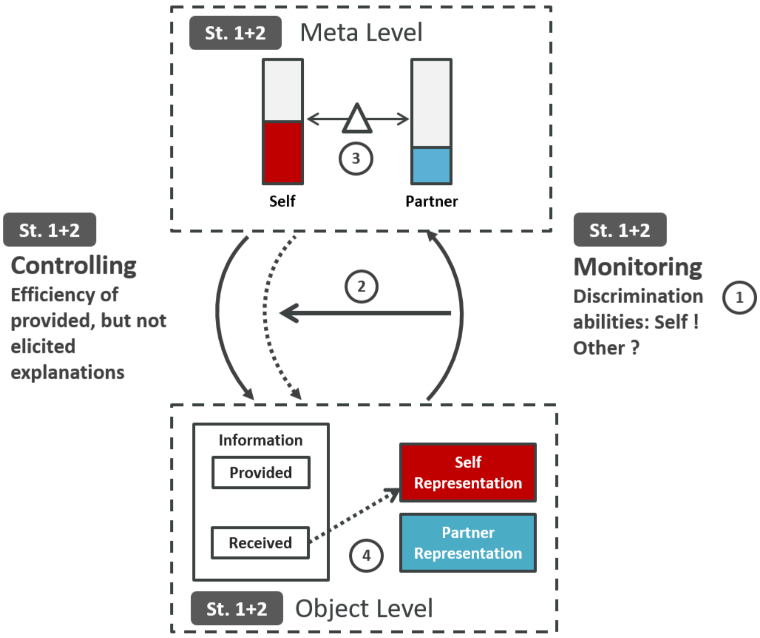

Figure 8. Results of the studies with regard to MEDIA focusing on absolute measurements 1. Like in individual metacognition, learners were able to monitor their knowledge item-by-item (relative partner monitoring was not examined so far) 2. The more accurate high-ability learners monitored themselves the more accurate they provided explanations for their partners, 3. To become an accurate explanation provider, high-ability learners needed to evaluate intragroup differences, 4. However, being more efficient in providing explanations did not help low-ability learners to gain more knowledge

However, monitoring own knowledge on item level in group learning settings was generally possible and about as (in)accurate as in individual learning scenarios (Maki, 1998). Next looking on the meta-level, Study 1 and Study 2 showed that learners are able to evaluate intragroup differences on item level, either when partner judgments were prompted (Study 1) or when learners identified well and poorly understood learning items of the partner during interaction (Study 2). But did the evaluation of intragroup differences influence learners control behavior, too? Even though in Study 1 no relationship could be found between relative monitoring accuracy and relative control accuracy of providing or receiving explanations, Study 2 could demonstrate that such a relationship indeed exists. More specifically, relative monitoring accuracy of high-ability learners was associated with their relative control accuracy to provide explanations. In contrast, learners only sparsely used questions for eliciting explanations, so that no relationship could be found between learners' relative monitoring accuracy and their relative control accuracy to ask questions. Examining the impact of accurate control behavior on the object-level, low-ability learners had no benefits from high-ability learners' efficiency to provide explanations (i.e. relative control accuracy to provide explanations), contrary to the findings from Thiede et al. (2003).

5.2 Strengths and Limitations

This dissertation developed an extension of Nelson and Narens' object-metal-level model for dyadic interaction. Having developed this extension can be already considered a strength on its own. Nonetheless, with this general approach come several limitations.

Even though the majority of cognitive states and processes of the proposed model could be examined by the presented experimental settings, some parts still need to be verified. The model suggests, for example, that learners hold own and partner goal states separately from each other. If learners in a dyad have different goals, different individual discrepancies might occur, although both learners show the same knowledge level. This could result in more complex patterns (or conflicts) of giving and receiving explanations, when one of the learners already reached her goals, while the other person still wants to gain more knowledge. Furthermore, the studies focused on the effect of relative monitoring accuracy of own knowledge to examine the influence on relative group control accuracy. Thus, I did not capture relative monitoring accuracy of partner knowledge. I will outline this aspect in more detail in the upcoming section "Implications for further research".

My research focused on the questions of whether learners represent own and partner state of *comprehension* or *knowledge* and how learners regulate their *understanding* of the learning material through collaborative interactions. Therefore, I reduced the scope of research to monitoring and controlling measurements, which focused on capturing comprehension, namely judgments of comprehension (JOC) and providing/eliciting explanations and elaborations. Other types of measurements, such as confidence judgments after an initial test (e.g., Huff & Nietfeld, 2009; Nietfeld et al., 2005; Schraw, 2009) or feelings of knowing (i.e. ability to identify the right answer from a range of options, once a first recall attempt was not successful; e.g., Koriat, 2000; Schwartz, 2008; Swerts & Krahmer, 2005), might reveal different results, as they tap different aspects of self- and partner-representation. For example, if learners perceive that partners are very confident in having correctly answered to a test item, whereas in fact the partner's answer was wrong, learners might invest more effort in convincing the partner of the correct answer, using more arguments and explanations. In the same way, if learners perceive that their partners have a high feeling of knowing, but are not able to directly recall the answer, learners might provide more hints than explanations to support their partners' knowledge recall. Additionally, in collaborative learning situations new types of metacognitive experiences might appear such as ease of communication, to which the results of the current dissertation might not be directly applicable.

With regard to external validity and the question of whether the findings are generalizable, readers should note that the presence and absence of evidence for elements of MEDIA came from controlled experimental studies, which were carried out in a more or less artificial lab situation. Thus, factors such as the environment, where learners meet to learn or work on their tasks, learners' psychological constitution (e.g. fatigue, stress), or the valence of the exam/task compared to other daily or educational tasks might influence learners' monitoring and control behavior in the group and should be analyzed in subsequent field studies. Even though students from different faculties participated in Studies 1 and 2, only university students were recruited. Moreover, Study 3 represented a quasi-experiment with a specific set of students from specific faculties. Hence, results cannot be generalized to students of all faculties and generally not to pupils of primary and secondary school. Study 1 did not capture actual behavior, but only intentions and expectations of learners' potential control behavior. While Study 2, captured learners' actual control behavior, this was only assessed in a computer-mediated chat environment. Consequently, Study 3 tried to provide a natural face-to-face environment for learners to interact. However, due to technical constraints group interactions could not be analyzed, so that findings of the distribution of provided and elicited explanations/elaborations, can only be generalized by referring to literature of small group learning (Saleh et al., 2005, 2007). Thus, future studies should always incorporate analyses of utterances to examine learners' natural group control behavior, and the impact of cues on learners' self- and partner-representations. Moreover, qualitative discourse analysis has to uncover how learners, especially low-ability learners, elicit explanations/elaborations from their partners.

Besides these limitations, I would argue that the present studies also carry some strengths. Jost et al. (1998) had complained that researchers from cognitive and social psychology worked independently in a field, which, by its referral to cognitive constructs, should be strongly related to each other. However, even after their review, not much work has been conducted to pick up and widen the field of social metacognition (for exceptions see Molinari et al., 2009; Sangin et al., 2011). The specific metacognitive model presented in this dissertation therefore serves as a new starting point to explicitly study collaborative learning dyads. For this, crucial ingredients of individual metacognition were translated into the field of group learning and tested in two different scenarios. Moreover, MEDIA provides the basis to study different factors, which might influence processes of monitoring or the exchange of information. For example, Study 2 and 3 discussed whether learners' experiences with such learning-test situations might have biased their judgments of own and partner knowledge at the end of the collaborative learning session, and suppressed cues such as the actual amount of provided or received information. But also other cognitive/motivational factors can be implemented into MEDIA and systematically tested, such as the influence of (dis)interest and accountability on the evaluation processes on the meta-level (Chaiken & Trope, 1999; Schiefele, 1991; Scholten et al., 2007). Also other factors, which can serve as cues for judgments, can be implemented and systematically tested in MEDIA, such as the influence of explanation quality on self- and partner-representations (Duschl, 2007; Toulmin, 1958).

So far, MEDIA concentrated on representations of knowledge and on knowledge exchange, but an advantage of the model is that it can be adapted to describe other metacognitive processes in the social context. For example, one might be interested in studying representations of attitudes, the monitoring of attitudes and investigating the information exchange needed to cause an attitude change (Ajzen & Fishbein, 1977; Eagly & Chaiken, 1993; Fazio, 2007). In a similar way, researchers could use the model to investigate emotional states, the monitoring of emotional states, and how people try to balance own and partner emotional states during conversation (Tamir, 2009; Zelazo & Cunningham, 2007). Hence, the model has the potential to mature into a general framework. However, much more research is needed until a generalization can be postulated. Together with the presented model, the dissertation provided the methodology to analyze the links between cognitive states, monitoring and social interactions by using the Actor-Partner-Interdependence-Model (APIM; Cook & Kenny, 2005; Kenny et al., 2006). The APIM is a flexible methodology that allows differentiating between actor and partner effects and taking the interdependence of the dyad into account. Just recently, the methodology has been extended to capture the interdependency in groups of more than two people (Garcia, Meagher, & Kenny, 2015). Therefore, the dissertation did not only provide a model to investigate social regulation of cooperative learning groups from a metacognitive perspective, but also provided the methodological tools to analyze the interplay between cognitive processes and social interactions.

This methodology was used to shed new light on the complex question which individual factors during collaboration (including interdependencies) might have an impact on (the evaluation of) own and partners' knowledge states. Therefore, the model gave further insight how and when cognitive growth appears in group learning by linking verbal utterances and cognitive activities. To further evaluate and validate the assumptions of MEDIA, the presented strengths and limitations already suggested potential research, which will be extended in the next section.

5.3 Implications for Future Research

In each of the three "Discussion" sections of Study 1-3, I already mentioned some implications, which were immediately related to the study and the presented results. In this section, I take a broader perspective and will try to outline implications for research focusing on cognitive aspects, on classroom interaction, and on technology enhanced group learning. The first part is therefore dedicated to researchers who are interested to validate or refute different elements in MEDIA or who are interested to study additional factors for their own work in the realm of MEDIA.

Implications for Research on MEDIA

Even though many elements of the model could be validated throughout the studies, one element could not be examined so far: the ability of the learner to discriminate between well and poorly understood learning items of the partner, named *relative partner-monitoring accuracy*. Relative partner-monitoring accuracy is an important factor during the collaborative learning phase, as the accuracy is, besides relative self-monitoring accuracy, the basis to evaluate intragroup differences for each learning item and to decide for which learning item further information needs to be exchanged. Only when relative self- *and* partner-monitoring accuracy is high, learners might enter an efficient information exchange. But the way how learners monitor their partners from one learning item to the next is relatively unclear. Swerts and colleagues did a number of studies to examine the amount of modalities needed (just audio recordings, just image/video recordings, or both) to judge the feeling of knowing of a target person accurately (Barkhuysen et al., 2005; Swerts & Krahmer, 2005). However, accuracy was determined by observers' abilities to meet the same judgment level as the recorded person herself. In the context of MEDIA and JOCs, it will be more interesting to know which factors lead observers to judge the appropriate knowledge level of the partner. MEDIA predicts that monitoring accuracy increases the more information the other person provides. Using a similar setting like in Swerts' studies, future work could alternate the amount of information provided by the recorded person for each learning item to validate this assumption. Furthermore, researchers might be interested what cognitive operations people actually perform to judge the information provided by the other person. With a high level of own knowledge, people might simply compare the provided information with their own understanding of the learning item, but when people are forced to judge the provided information without any prior knowledge to the learning item, they might need to examine the information according to quality criteria. Therefore, researchers could alternate the amount of prior knowledge of the observer or the quality of provided information of the target person to explore which factor enables observers to evaluate the presented information of the target person.

But MEDIA not only predicts that the received information from the partner has an influence on partner-monitoring accuracy, but also the information the learner provides for her partner. From *Audience Design* it is known that people adapt their speech depending whether they talk with a close friend or with a complete stranger (Clark & Murphy, 1982; Horton & Gerrig, 2002). Also when students were asked to write down their understanding of a topic for themselves or for another student, they significantly changed their style of writing and used more figurative speech for example (Fussell & Krauss, 1989). In the context of social metacognition researchers might be interested in examining if students will also change their partner-judgments once they adapted their speech by providing for instance more or less information for partners with low or high prior knowledge. Furthermore, personality factors such as need for cognitive closure (Richter & Kruglanski, 1999; Webster & Kruglanski, 1997) might influence the way learners respond to their partners and how they might judge partner knowledge afterwards. In contrast to settings in which learners simply have to judge a person's knowledge by the amount of received information (see the studies of Swerts and colleagues in the previous paragraph), learners' adaptations of speech to a partner's characteristics and their influence on partner judgments can only be examined with a real learning partner during the course of a collaborative learning session.

Once researchers deeply study the interaction of real dyads, the interplay between self and partner-monitoring might be of interest, too. As mentioned above, relative self- and partner-monitoring accuracy should be high in order to precisely evaluate intragroup differences and to engage in efficient information exchange. Therefore, future research should investigate intermediate stages of judging own and partner's learning to analyze how both accuracy values change during a collaborative learning phase. As people showed notoriously low self-monitoring accuracy in discriminating well and poorly understood learning items, they might show low partner-monitoring accuracy, too. Thus, factors of how relative partner-monitoring accuracy can be increased could be transferred from studies of individual self-monitoring accuracy and adapted to the purposes of the collaborative context. For instance, research on individual metacognition showed benefits via the so-called generating effect (e.g. writing down keywords or verbally summarizing the information everyone in the group mentioned concerning each learning item before making the respective judgments, Dunlosky & Lipko, 2007; Thiede et al., 2003). Moreover, with intermediate stages of capturing learners' monitoring abilities researchers can also analyze learners' intermediate regulation abilities within a collaborative learning phase. Learners might show relatively low group control accuracy at the beginning of a collaborative learning session, but the more learners interact with each other group control accuracy might increase together with relative self- and partner-monitoring accuracy. Together, MEDIA urges cognitive scientists and psychologists, who are interested in cooperative learning, to examine more deeply the process of collaboration and to capture the different metacognitive aspects. Besides the immediate question how to uncover a better understanding of the different metacognitive aspects of monitoring and controlling, further questions appear in the realm of classroom practice.

Implications for the Classroom

With regard to the formation of dyads during classroom time, the presented studies focused on *heterogeneous* dyads (Lou et al., 1996, 2001; Saleh et al., 2005, 2007; Webb, 1989, 1991), but researchers might be equally interested to investigate the regulatory mechanisms of *homogeneous* dyads. Especially average- or high-ability learners show equally good learning potentials. However, group regulation within such group compositions might be a difficult issue. When learners gain new knowledge equally fast through the course of collaboration assumed similarity can play the role of a consensus facilitator. But if learners develop differently within a dyad, for instance due to motivation or personal interest (Krapp, 1999; Schiefele & Krapp, 1996; Schiefele, 1991), assumed similarity will bias learners towards a false consensus. Thus, researchers should investigate more flexible and adaptive composition techniques, which take learners' learning pace into account. Moreover, placing learners from homogeneous to heterogeneous groups and backwards might increase learners' sensitivity to different partner knowledge levels and monitoring abilities of own knowledge levels.

The present dissertation also tapped into a rather neglected field in educational research: alternating levels of task difficulty. One alternative explanation for the obtained results in Study 2 and 3 concerning absolute self- and partner monitoring accuracy was that learners' monitoring (in)accuracy might have been caused by a tendency to judge oneself and the partner towards a socially desired performance value (Kruger & Dunning, 1999), which can be termed *false average bias*. According to the results, learners in the presented experiments were not only ignoring their provided or received explanations, but also the difficulty or ease of the task as potential cues. The desired performance value often stems from prior experiences with similar learning settings and from socially accepted mean values, acquired over the years of a learner's educational life. Therefore, researchers should study more deeply how self- and partner-monitoring accuracy change in the context of alternating task difficulties, and in the context of alternating test settings, in order to develop monitoring techniques, which might counteract the false average bias.

Speaking of techniques and trainings, educators might want to support their students in monitoring self- and partner knowledge. Previous research has shown that people can be primed to be more sensitive to partners' perspectives (Galinsky, Maddux, Gilin, & White, 2008; Gockel & Brauner, 2013). Perspective taking techniques try to emphasize the viewpoints of other persons by priming learners with tasks, which requires them to put themselves into the shoes of the other person. This way, Gockel and Brauner (2013) could show that participants being primed to take the perspective of another person were subsequently more accurate in judging partners' expertise in a decision making task, regardless of whether the team held a distributed or integrated knowledge about the problem. The results of the current dissertation have shown that learners need to be aware of both own *and* partner perspective when it comes to monitoring and evaluating the respective knowledge state. Therefore, future techniques for social metacognition should incorporate tasks which force learners to take own *and* other persons' perspectives into account and to detect particular differences. This kind of difference recognition has been termed detection of anomalies by comparing different meanings of an issue and can be transferred to the social context of learning groups, for which different states of comprehension need to be compared (Graesser & McMahen, 1993). Comparably, the technique might prompt learners to look for hidden aspects of own and partner viewpoints, or intersections of both learners' expertise, in a more cooperative scenario. Once an anomaly is detected, learners might question the partner if she recognized the difference in understanding, too, so that better collaborative elaborations occur, leading to more efficient information exchange. Of course, the priming technique can be extended with elements in a way that emphasizes specific control behaviors, so that the whole monitoring and control cycle will be part of the training. Thus, learners might not only be primed to put themselves into the shoes of another person, but also to identify differences for example in size and color as well as to create solutions so that everyone can wear the shoes of the other person. As was shown, MEDIA helps educational researchers to differentiate between distinct metacognitive activities during collaboration so that specific interventions can be developed, which either counteract biases such as false consensus and false average beliefs, or that teach students necessary metacognitive skills to better monitor and control the collaborative learning phase. However, teachers might not always be around to instruct learners to become aware of different group constellations, and especially of differences in knowledge states. Advances in technology could help to overcome this problem.

Implications for Technology Enhanced Group Learning

According to the main findings, learners have troublesome difficulties monitoring own and partner knowledge accurately, while their evaluation of partner knowledge is highly influenced by their own knowledge and understanding of the learning material. Enabling learners to better monitor and differentiate between own and partner knowledge is therefore the key to an efficient knowledge exchange, and technology can play different roles to guide learners towards this goal. First, technical environments can use metacognitive prompts which urge learners in pre-defined time intervals to engage in collaborative activities such as reminding learners to state their full understanding of the learning material. Such prompts might allow learners to better monitor their own and partner knowledge. Additionally, not only monitoring but also evaluation prompts can be implemented to remind learners, for instance, to regularly evaluate intragroup differences for each learning item (compare Stadtler & Bromme, 2008). Therefore, technology can play the role of a trigger activating the desired metacognitive processes. Second, computers can help identifying complex knowledge structures, which is often the case in problem-based cooperative scenarios, and which cannot be clearly distinguished into separated learning items. Thus, digital visualization tools such as mind mapping tools can help learners to gain a better picture about own and partner knowledge structures by arranging different information elements in a meaningful way and eliciting links among these elements (e.g., Novak & Gowin, 1984). While visualizing their knowledge, learners are asked to externalize and distribute their cognition into digital artifacts (Hollan, Hutchins, & Kirsh, 2000; Hutchins, 1995). Therefore, learners might monitor own and partner knowledge more accurately (compare generation effect), and create a basis to better identify differences and anomalies in their knowledge structures, which should lead to a more effective knowledge exchange.

Third, computers can be connected and exchange information, which offers the possibility of capturing and exchanging each learner's (meta)cognitive states and activities. For example, learners can be provided with partner' self-judgment when it is transferred to a learner's own device (compare Study 2). Although partner' self-judgments might be not highly accurate, it might be more accurate than learners making partner-judgments by themselves due to the false consensus bias. Moreover, combining interconnectivity and visualization tools has the potential to create states of *group knowledge awareness*, which proved its usefulness throughout numerous experiments in different collaborative learning scenarios (Dehler et al., 2011; Engelmann & Hesse, 2010; Sangin et al., 2011). So far I outlined how technological environments can prompt, capture and visualize *subjective* monitoring activities of learners to improve the evaluation of knowledge states and provide guidance for collaborative interactions. However, computers can also assess learners' *objective* knowledge state in a regular test-feedback cycle. While in *formative assessment* usually teachers take over the role of feedback-providers (Nicol & Macfarlane-Dick, 2006), learners can also be provided with an analysis of their learning performance by technological environments which have been termed *adaptive feedback systems* or *intelligent tutoring systems* (Ma, Adesope, Nesbit, & Liu, 2014; Romero & Ventura, 2007; van Lehn, 2011; Walker, Rummel, & Koedinger, 2009). These systems are typically tailored to individual learners, and adapt the learning material to the actual knowledge needs of each individual learner. Throughout the course of collaboration, computers could analyze and track both learners' knowledge gains in regular time intervals, compare and evaluate learners' knowledge states item-by-item, and feed back these analyses to the learners. Depending on the degree of adaptation and tutoring, the feedback can simply prompt learner knowledge states, or the feedback might involve specific instructions for which learning items information needs to be exchanged. In either way, having an objective assessment of both own and partner knowledge state at hand, learners are not required to concentrate on an accurate monitoring, but can directly focus on intragroup differences and subsequent group regulation, which should improve the knowledge exchange efficiency.

As seen, the presented model in this dissertation provides cognitive psychologists, as well as educational scientists the basis to decide at which point of the socio-metacognitive process during cooperative group learning they want to locate their research. For this, researchers need to clarify whether they want to explore cognitive/motivational or environmental factors, foster learners' social regulatory skills, or implement computational aids that help learners to identify and regulate differences in knowledge and expertise. For this, MEDIA provides the starting point with clearly specified dependencies among different cognitive states and processes.

5.4 Conclusion

Learning has changed substantially in the last fifty years. Learning in groups has become part of students' daily educational life, whether they prepare collaboratively for the next exam or whether they have to coordinate their team work for a problem-solving task. Additionally, teachers employ cooperative work in the classroom even more since educational institutions in secondary and post-secondary education have the technical infrastructure to implement *blended learning* strategies such as the *inverted classroom* (Bonk & Graham, 2006; Strayer, 2012).

Within all of these collaborative scenarios, group regulation is a mandatory skill of every learner, and MEDIA was developed to explain how individual group members monitor mental states of knowledge/expertise and how knowledge exchange should be controlled within the learning group. Throughout the empirical chapters, the dissertation elicited various elements of MEDIA, such as self- and partner-representations of knowledge, the monitoring of self- and partner-representations, the evaluation of individual discrepancies and intragroup differences, and the control behavior including providing and receiving explanations. Further, the results point to problems during monitoring and controlling such as strong biases during judgments of partner knowledge or missing benefits of efficient knowledge exchange. Therefore, more work has to be done identifying further influencing factors and possible countermeasures, so that learners in cooperative learning can accurately monitor and control their actions. Together, the dissertation brings research in social metacognition forward and encourages scholars from fields of psychology and educational sciences to investigate more deeply the regulatory mechanisms of groups from an individual cognitive perspective so that students receive the support they need to become self-determined and accountable collaborators.

References

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*(5), 888–918.

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. http://doi.org/10.1037/a0015928

Aronson, E. (1997). *The jigsaw classroom: Building cooperation in the classroom*. Scott Foresman & Company.

Astington, J. W., & Jenkins, J. M. (1995). Theory of mind development and social understanding. *Cognition & Emotion*, *9*(2-3), 151–165. http://doi.org/10.1080/02699939508409006

Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.

Baker, M., Andriessen, J., Lund, K., van Amelsvoort, M., & Quignard, M. (2007). Rainbow: A framework for analysing computer-mediated pedagogical debates. *International Journal of Computer-Supported Collaborative Learning*, *2*(2-3), 315–357. http://doi.org/10.1007/s11412-007-9022-4

Baker, M., Hansen, T., & Joiner, R. (1999). The role of grounding in collaborative problem solving tasks. In P. Dillenbourg (Ed.), *Collaborative Learning : Cognitive and Computational Approaches* (1st ed, pp. 31–64). Amsterdam ; New York: Elsevier Science.

Barkhuysen, P., Krahmer, E., & Swerts, M. (2005). Problem detection in human–machine interactions based on facial expressions of users. *Speech Communication*, *45*(3), 343–359. http://doi.org/10.1016/j.specom.2004.10.004

Barron, B. J. S., Schwartz, D. L., Vye, N. J., Moore, A., Petrosino, A., Zech, L., & Bransford, J. D. (1998). Doing with understanding: Lessons from research on problem- and project-based learning. *Journal of the Learning Sciences*, *7*(3-4), 271–311. http://doi.org/10.1080/10508406.1998.9672056

Bayrhuber, H., Kull, U., & Linder, H. (2005). *Linder Biologie [Linder Biology]* (22nd ed.). Braunschweig: Schroedel.

Bertucci, A., Johnson, D. W., Johnson, R. T., & Conte, S. (2011). The effects of task and resource interdependence on achievement and social support: an exploratory study of Italian children. *The Journal of Psychology*, *145*(4), 343–360.

Bertucci, A., Johnson, D. W., Johnson, R. T., & Conte, S. (2012). Influence of group processing on achievement and perception of social and academic support in elementary inexperienced cooperative learning groups. *The Journal of Educational Research*, *105*(5), 329–335.

Biesanz, J. C., West, S. G., & Millevoi, A. (2007). What do you learn about someone over time? The relationship between length of acquaintance and consensus and self-other agreement in judgments of personality. *Journal of Personality and Social Psychology*, *92*(1), 119–35.

Birch, S. A. J., & Bloom, P. (2004). Understanding children's and adults' limitations in mental state reasoning. *Trends in Cognitive Sciences*, *8*(6), 255–260. http://doi.org/10.1016/j.tics.2004.04.011

Bishop, L. J., & Verleger, M. (2013). The flipped classroom : A survey of the research. *Proccedings of the Annual Conference of the American Society for Engineering Education*, 6219. http://doi.org/10.1109/FIE.2013.6684807

Bonk, C. J., & Graham, C. R. The handbook of blended learning, In 3–21 (2006). http://doi.org/Book Review

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education*, *53*(4), 1207–1217. http://doi.org/10.1016/j.compedu.2009.06.004

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). *The psychology of learning and motivation: Advances in research and theory*. *Psychology of Learning and Motivation* (Vol. 53). Elsevier. http://doi.org/10.1016/S0079-7421(10)53008-1

Brünken, R., Steinbacher, S., Plass, J. L., & Leutner, D. (2002). Assessment of cognitive load in multimedia learning using dual-task methodology. *Experimental Psychology*, *49*(2), 109–119.

Buchs, C., Butera, F., & Mugny, G. (2004). Resource interdependence, student interactions and performance in cooperative learning. *Educational Psychology*, *24*(3), 291–314.

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.

Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, *27*(1), 79–96. http://doi.org/10.1017/S0140525X04000032

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*(1), 19–35. http://doi.org/10.1037/0033-295X.97.1.19

Chaiken, S., & Trope, Y. (1999). Dual-process theories in social psychology. In *Social Psychology* (p. 676).

Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. *Advances in Psychology*, *9*(C), 287–299.

Cohen, E. G. (1994). Restructuring the classroom: Conditions for productive small groups. *Review of Educational Research*, *64*(1), 1–35.

Cook, W. L., & Kenny, D. A. (2005). The Actor-Partner Interdependence Model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development*, *29*(2), 101–109. http://doi.org/10.1080/01650250444000405

Dansereau, D. F. (1988). Cooperative learning strategies. In C. E. Weinstein, E. T. Goetz, & P. A. Alexander (Eds.), *Learning and study strategies: Issues in assessment, instruction, and evaluation.* (pp. 103–120). Academic Press.

de Bruin, A. B. H., Thiede, K. W., Camp, G., & Redford, J. S. (2011). Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology*, *109*(3), 294–310. http://doi.org/10.1016/j.jecp.2011.02.005

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.

Dinet, J., & Vivian, R. (2012). The impact of friendship on synchronous collaborative retrieval tasks in the primary school. *British Journal of Educational Technology*, *43*(3), 439–447.

Dunlosky, J., & Hertzog, C. (1998). Training programs to improve learning in later adulthood: Helping older adults educate themselves. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 249–275).

Dunlosky, J., & Lipko, A. R. (2007). Metacomprehension: A brief history and how to improve its accuracy. *Current Directions in Psychological Science*, *16*(4), 228–232. http://doi.org/10.1111/j.1467-8721.2007.00509.x

Dunn, M., Thomas, J. O., Swift, W., & Burns, L. (2012). Elite athletes' estimates of the prevalence of illicit drug use: Evidence for the false consensus effect. *Drug and Alcohol Review*, *31*(1), 27–32.

Duschl, R. A. (2007). Quality argumentation and epistemic criteria. In S. Erduran & M. P. Jiménez-Aleixandre (Eds.), *Argumentaion in Science Education* (pp. 159–175). Springer Netherlands. http://doi.org/10.1007/978-1-4020-6670-2_8

Eagly, A. H., & Chaiken, S. (1993). The nature of attitudes. In *The Psychology of Attitudes* (pp. 1–21). Harcourt Brace Jovanovich College Publishers.

Engelmann, T., Dehler, J., Bodemer, D., & Buder, J. (2009). Knowledge awareness in CSCL: A psychological perspective. *Computers in Human Behavior*, *25*(4), 949–960. http://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. http://doi.org/10.1007/s11412-010-9089-1

Fazio, R. H. (2007). Attitudes as object–evaluation associations of varying strength. *Social Cognition*, *25*(5), 603–637.

Fiske, S. T. (1993). Social cognition and social perception. *Annual Review of Psychology*, *44*(1), 155–194. http://doi.org/10.1016/j.tics.2006.11.005

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

Foley, C., & Smeaton, A. F. (2009). Evaluation of coordination techniques in synchronous collaborative information retrieval. In *Proceeding of the Joint conference on Digital libraries - (JCDL 08) Workshop on Collaborative Information Retrieval* (pp. 1–4).

Fussell, S. R., & Krauss, R. M. (1989). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology*, *25*(3), 203–219. http://doi.org/10.1016/0022-1031(89)90019-X

Galinsky, A. D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent in negotiations. *Psychological Science*, *19*(4), 378–384. http://doi.org/10.1111/j.1467-9280.2008.02096.x

Garcia, R. L., Meagher, B. R., & Kenny, D. A. (2015). Analyzing the effects of group members' characteristics: A guide to the group actor-partner interdependence model. *Group Processes & Intergroup Relations*, *18*(3), 315–328. http://doi.org/10.1177/1368430214556370

Gerjets, P., & Hellenthal-Schorr, T. (2008). Competent information search in the World Wide Web: Development and evaluation of a web training for pupils. *Computers in Human Behavior*, *24*(3), 693–715.

Gibbons, F. X., & Buunk, B. P. (1999). Individual differences in social comparison: Development of a scale of social comparison orientation. *Journal of Personality and Social Psychology*. http://doi.org/10.1037/0022-3514.76.1.129

Gijlers, H., & De Jong, T. (2005). The relation between prior knowledge and students' collaborative discovery learning processes. *Journal of Research in Science Teaching*, *42*(3), 264–282.

Gockel, C., & Brauner, E. (2013). The benefits of stepping into others' shoes: Perspective taking strengthens transactive memory. *Basic and Applied Social Psychology*, *35*(2), 222–230. http://doi.org/10.1080/01973533.2013.764303

Golub, M., & Buchs, C. (2014). Preparing pupils to cooperate during cooperative controversy in grade 6: a way to increase positive interactions and learning? *European Journal of Psychology of Education*, pp. 1–14.

Gonzalez, R., & Nelson, T. O. (1996). Measuring ordinal association in situations that contain tied scores. *Psychological Bulletin*, *119*(1), 159–165.

Goos, M., Galbraith, P., & Renshaw, P. (2002). Socially mediated metacognition: creating collaborative zones of proximal development in small group problem solving. *Educational Studies in Mathematics*, *49*(2), 193–223.

Graesser, A. C., & McMahen, C. L. (1993). Anomalous information triggers questions when adults solve quantitative problems and comprehend stories. *Journal of Educational Psychology*, *85*(1), 136–151.

Gureckis, T. M., & Markant, D. B. (2012). Self-Directed Learning: A cognitive and computational perspective. *Perspectives on Psychological Science*, *7*(5), 464–481.

Hacker, D. J. (1998). Definitions and empirical foundations. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 1–23).

Hadwin, A. F., Järvelä, S., & Miller, M. (2011). Self-regulated, co-regulated, and socially shared regulation of learning. In *Handbook of self-regulation of learning and performance* (pp. 65–84).

Hilsch, P. (2012). *Das Mittelalter - die Epoche*. UTB. Retrieved from http://books.google.com/books?id=ccQ7Jgtg1l0C&pgis=1

Hollan, J. D., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, *7*(2), 174–196. http://doi.org/10.1145/353485.353487

Horton, W. S., & Gerrig, R. J. (2002). Speakers' experiences and audience design: Knowing when and knowing how to adjust utterances to addressees. *Journal of Memory and Language*, *47*(4), 589–606.

Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning*, *4*(2), 161–176.

Hummel, H. G. K., Paas, F., & Koper, E. J. R. (2004). Cueing for transfer in multimedia programmes: Process worksheets vs. worked-out examples. *Journal of Computer Assisted Learning*, *20*(5), 387–397.

Hutchins, E. (1995). Cognition in the wild. *MIT Press*. http://doi.org/10.1023/A:1008642111457

Hyldegard, J. (2006). Collaborative information behaviour––exploring Kuhlthau. *Information Processing & Management*, *42*(1), 276–298.

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. http://doi.org/10.1016/j.learninstruc.2010.05.002

Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: Elements of a cognitive IR theory. *Journal of Documentation*, *52*(1), 3–50.

John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: the big five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, *61*(4), 521–551.

Johnson, D. W., & Johnson, R. T. (1999). Making cooperative learning work. *Theory Into Practice*, *38*(2), 67–73. http://doi.org/10.1080/00405849909543834

Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review*, *19*(1), 15–29. http://doi.org/10.1007/s10648-006-9038-8

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (1989). Impact of goal and resource interdependence on problem-solving success. *The Journal of Social Psychology*.

Jonas, K. J., & Huguet, P. (2008). What day is today? A social-psychological investigation into the process of time orientation. *Personality and Social Psychology Bulletin*, *34*(3), 353–365. http://doi.org/10.1177/0146167207311202

Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, *48*(4), 63–85.

Jost, J. T., Kruglanski, A. W., & Nelson, T. O. (1998). Social metacognition: An expansionist review. *Personality and Social Psychology Review*, *2*(2), 137–154. http://doi.org/10.1207/s15327957pspr0202_6

Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language*, *35*(2), 157–175.

Kelly, J. R., & Karau, S. J. (1999). Group decision making: The effects of initial preferences and time pressure. *Personality and Social Psychology Bulletin*, *25*(11), 1342–1354. http://doi.org/10.1177/0146167299259002

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. Guilford Press.

Kenny, D. A., & West, T. V. (2010). Similarity and agreement in self-and other perception: a meta-analysis. *Personality and Social Psychology Review*, *14*(2), 196–213.

Khosa, D. K., & Volet, S. E. (2014). Productive group engagement in cognitive activity and metacognitive regulation during collaborative learning: can it explain differences in students' conceptual understanding? *Metacognition and Learning*, 287–307. http://doi.org/10.1007/s11409-014-9117-z

Kluwe, R. H. (1982). Cognitive Knowledge and Executive Control: Metacognition. *Animal Mind – Human Mind*, *21*, 201–224. http://doi.org/10.1007/978-3-642-68469-2_12

Korchmaros, J. D., & Kenny, D. A. (2006). An evolutionary and close-relationship model of helping. *Journal of Social and Personal Relationships*, *23*(1), 21–43. http://doi.org/10.1177/0265407506060176

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*(4), 349–370. http://doi.org/10.1037/0096-3445.126.4.349

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, *9*(2), 149–171.

Koriat, A. (2012). The relationships between monitoring, regulation and performance. *Learning and Instruction*, *22*(4), 296–298.

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.

Körner, A., Geyer, M., Roth, M., Drapeau, M., Schmutzer, G., Albani, C., ... Brähler, E. (2008). Personality assessment with the NEO-five-factor inventory: The 30-item-short-version (NEO-FFI-30). *Psychotherapie, Psychosomatik, medizinische Psychologie*, *58*(6), 238–245. http://doi.org/10.1055/s-2007-986199

Krapp, A. (1999). Interest, motivation and learning: An educational-psychological perspective. *European Journal of Psychology of Education*, *14*(1), 23–40. http://doi.org/10.1007/BF03173109

Krueger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology*, *67*(4), 596–610. http://doi.org/10.1037/0022-3514.68.4.579

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134.

Kulthau, C. C. (1991). Inside the search process: Information Seeking from the user's perspective. *Journal of the American Society for Information Science and Technology*, *42*(5), 361–371.

Ladegaard, H. J. (1995). Audience design revisited: Persons, roles and power relations in speech interactions. *Language and Communication*, *15*(1), 89–101.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. http://doi.org/10.2307/2529310

Lazonder, A. W. (2005). Do two heads search better than one? Effects of student collaboration on web search behaviour and search outcomes. *British Journal of Educational Technology*, *36*(3), 465–475.

Leary, M. R., & Kowalski, R. M. (1990). Impression management : A literature review and two-component model. *Psychological Bulletin*, *107*(I), 34–47.

Leonard, J. (2001). How group composition influenced the achievement of sixth-grade mathematics students. *Mathematical Thinking and Learning*, *3*(2-3), 175–200. http://doi.org/10.1080/10986065.2001.9679972

Letzring, T. D., Wells, S. M., & Funder, D. C. (2006). Information quantity and quality affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology*, *91*(1), 111–123.

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & Apollonia, S. (1996). Within-class grouping : A meta-analysis. *Review of Educational Research*, *66*(4), 423–458. http://doi.org/10.3102/00346543066004423

Lou, Y., Abrami, P. C., Spence, J. C., Poulsen, C., Chambers, B., & D'Apollonia, S. (2001). Small group and individual learning with technology: A meta-analysis. *Review of Educational Research*, *71*(3), 449–521. http://doi.org/10.3102/00346543071003449

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, *106*(4), 1–18.

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

Marchionini, G. (1995). Information-seeking perspective and framework (Chapter 3). In *Information seeking in electronic environments (Cambridge series on human computer interaction: Vol.9)* (pp. 27–60). Cambridge University Press. http://doi.org/10.1017/CBO9780511626388

Marks, G., & Miller, N. (1987). Ten years of research on the false-consensus effect: An empirical and theoretical review. *Psychological Bulletin*, *102*(1), 72–90.

Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. *The Cambridge Handbook of Multimedia Learning*, 31–48. http://doi.org/10.1207/s15326985ep4102_2

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science*, *18*(3), 159–163. http://doi.org/10.1111/j.1467-8721.2009.01628.x

Molinari, G., Sangin, M., Dillenbourg, P., & Nüssli, M.-A. (2009). Knowledge interdependence with the partner, accuracy of mutual knowledge model and computer-supported collaborative learning. *European Journal of Psychology of Education*, *24*(2), 129–144. http://doi.org/10.1007/BF03173006

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, *115*(2), 502–517. http://doi.org/10.1037/0033-295X.115.2.502

Moreno, R., & Mayer, R. (2007). Interactive Multimodal Learning Environments. *Educational Psychology Review*, *19*(3), 309–326. http://doi.org/10.1007/s10648-007-9047-2

Morris, M. R., Lombardo, J., & Wigdor, D. (2010). WeSearch: Supporting collaborative search and sensemaking on a tabletop display. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 401–410). ACM.

Morris, M. R., Paepcke, A., & Winograd, T. (2006). TeamSearch: Comparing techniques for co-present collaborative search of digital media. In *Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems, TABLETOP'06* (Vol. 2006, pp. 97–104).

Nam, C. W., & Zellner, R. D. (2011). The relative effects of positive interdependence and group processing on student achievement and attitude in online cooperative learning. *Computers & Education*, *56*(3), 680–688. http://doi.org/10.1016/j.compedu.2010.10.010

Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In *Metacognition: Knowing about knowing* (pp. 1–25).

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.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. http://doi.org/10.1080/03075070600572090

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2005). Metacognitive monitoring accuracy and student performance in the postsecondary classroom. *The Journal of Experimental Educational*, *74*(1), 7–28. http://doi.org/10.2307/20157410

Nietfeld, J. L., & Schraw, G. (2002). The effect of knowledge and strategy training on monitoring accuracy. *The Journal of Educational Research*, *95*(3), 131–142. http://doi.org/10.1080/00220670209596583

Novak, J. D., & Gowin, B. D. (1984). *Learning how to learn.* London: Cambridge University Press.

O'Donnell, A. M. (1996). Effects of explicit incentives on scripted and unscripted cooperation. *Journal of Educational Psychology*, *88*(1), 74–86.

O'Donnell, A. M., & O'Kelly, J. (1994). Learning from peers: Beyond the rhetoric of positive results. *Educational Psychology Review*, *6*(4), 321–349.

Ortiz, A. E., Johnson, D. W., & Johnson, R. T. (1996). The effect of positive goal and resource interdependence on individual performance. *The Journal of Social Psychology*, *136*(2), 243–249.

Overall, N. C., Fletcher, G. J. O., & Kenny, D. A. (2012). When bias and insecurity promote accuracy: Mean-level bias and tracking accuracy in couples' conflict discussions. *Personality and Social Psychology Bulletin*, *38*(5), 642–655.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, *1*(2), 117–175.

Palincsar, A. S., Brown, A. L., & Martin, S. M. (1987). Peer interaction in reading comprehension instruction. *Educational Psychologist*.

Pickens, J., Golovchinsky, G., Shah, C., Qvarfordt, P., & Back, M. (2008). Algorithmic mediation for collaborative exploratory search. *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '08*, 315.

Rawson, K. A., Dunlosky, J., & Thiede, K. W. (2000). The rereading effect: Metacomprehension accuracy improves across reading trials. *Memory & Cognition*, *28*(6), 1004–1010. http://doi.org/10.3758/BF03209348

Ray, D. G., Neugebauer, J., Sassenberg, K., Buder, J., & Hesse, F. W. (2012). Motivated shortcomings in explanation: The role of comparative self-evaluation and awareness of explanation recipient's knowledge. *Journal of Experimental Psychology: General*, *142*(2), 445–457. http://doi.org/10.1037/a0029339

Reder, L. M. (1980). *The role of elaboration in the comprehension and retention of prose: A critical review*. *Review of Educational Research* (Vol. 50).

Reis, H. T., Maniaci, M. R., Caprariello, P. a, Eastwick, P. W., & Finkel, E. J. (2011). Familiarity does indeed promote attraction in live interaction. *Journal of Personality and Social Psychology*, *101*(3), 557–70.

Rhodes, M. G., & Tauber, S. K. (2011). The influence of delaying judgments of learning on metacognitive accuracy: A meta-analytic review. *Psychological Bulletin*, *137*(1), 131–148. http://doi.org/10.1037/a0021705

Richter, L., & Kruglanski, a. W. (1999). Motivated Search for Common Ground: Need for Closure Effects on Audience Design in Interpersonal Communication. *Personality and Social Psychology Bulletin*, *25*(9), 1101–1114.

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, *33*(1), 135–146. http://doi.org/10.1016/j.eswa.2006.04.005

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": Egocentric bias in social-perception and attribution processes. *Journal of Experimental Social Psychology*, *13*(3), 279–301.

Saleh, M., Lazonder, A. W., & de Jong, T. (2005). Effects of within-class ability grouping on social interaction, achievement, and motivation. *Instructional Science*, *33*(2), 105–119. http://doi.org/10.1007/s11251-004-6405-z

Saleh, M., Lazonder, A. W., & de Jong, T. (2007). Structuring collaboration in mixed-ability groups to promote verbal interaction, learning, and motivation of average-ability students. *Contemporary Educational Psychology*, *32*(3), 314–331. http://doi.org/10.1016/j.cedpsych.2006.05.001

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. http://doi.org/10.1016/j.chb.2010.05.032

Schiefele, U. (1991). Interest, Learning, and Motivation. *Educational Psychologist*, *26*(3-4), 299–323. http://doi.org/10.1080/00461520.1991.9653136

Schiefele, U., & Krapp, A. (1996). Topic interest and free recall of expository text. *Learning and Individual Differences*, *8*(2), 141–160. http://doi.org/10.1016/S1041-6080(96)90030-8

Scholten, L., van Knippenberg, D., Nijstad, B. A., & De Dreu, C. K. W. (2007). Motivated information processing and group decision-making: Effects of process accountability on information processing and decision quality. *Journal of Experimental Social Psychology*, *43*(4), 539–552. http://doi.org/10.1016/j.jesp.2006.05.010

Schraw, G. (2009a). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and Learning*, *4*(1), 33–45.

Schraw, G. (2009b). Measuring metacognitive judgments. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (p. 415). New York.

Schwartz, B. L. (2008). Working memory load differentially affects tip-of-the-tongue states and feeling-of-knowing judgments. *Memory & Cognition*, *36*(1), 9–19.

Shah, C. (2014). Collaborative information seeking. *Journal of the Association for Information Science and Technology*, *65*(2), 215–236. http://doi.org/10.1002/asi.22977

Slavin, R. (1983). When does cooperative learning increase student achievement? *Psychological Bulletin*, *94*(3), 429–445.

Slavin, R. (2011). Instruction based on cooperative learning. *Handbook of Research on Learning and Instruction*, 344–360.

Soderstrom, N. C., & McCabe, D. P. (2011). The interplay between value and relatedness as bases for metacognitive monitoring and control: Evidence for agenda-based monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1236–1242. http://doi.org/10.1037/a0023548

Spörer, N., Brunstein, J. C., & Kieschke, U. (2009). Improving students' reading comprehension skills: Effects of strategy instruction and reciprocal teaching. *Learning and Instruction*, *19*(3), 272–286.

Stadtler, M., & Bromme, R. (2008). Effects of the metacognitive computer-tool met.a.ware on the web search of laypersons. *Computers in Human Behavior*, *24*(3), 716–737.

Stasser, G., & Titus, W. (2003). Hidden profiles: A brief history. *Psychological Inquiry*, *14*(3), 304–313.

Stone, B. B. (2012). Flip your classroom to increase active learning and student engagement. *28th Annual Conference on Distance Teaching & Learning*, 1–5.

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research*, *15*(2), 171–193.

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

Swerts, M., & Krahmer, E. (2005). Audiovisual prosody and feeling of knowing. *Journal of Memory and Language*, *53*(1), 81–94. http://doi.org/10.1016/j.jml.2005.02.003

Tamir, M. (2009). What do people want to feel and why? Pleasure and utility in emotion regulation. *Current Directions in Psychological Science*, *18*(2), 101–105. http://doi.org/10.1111/j.1467-8721.2009.01617.x

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. http://doi.org/10.1037/0022-0663.95.1.66

Thiede, K. W., Griffin, T. D., Wiley, J., & Redford, J. S. (2009). Metacognitive monitoring during and after reading. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education.* (pp. 85–106). New York: Routledge.

Toulmin, S. E. (1958). *The uses of argument* (2003rd ed.). Cambridge: Cambridge University Press.

van der Laan Smith, J., & Spindle, R. M. (2007). The impact of group formation in a cooperative learning environment. *Journal of Accounting Education*, *25*(4), 153–167. http://doi.org/10.1016/j.jaccedu.2007.09.002

van Lehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221. http://doi.org/10.1080/00461520.2011.611369

van Merriënboer, J. J. G., & Kester, L. (2008). Whole-task models in education. *Handbook of Research on Educational Communications and Technology*, 441–457.

Vazire, S. (2010). Who knows what about a person? The self-other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, *98*(2), 281–300.

Walker, E., Rummel, N., & Koedinger, K. R. (2009). CTRL: A research framework for providing adaptive collaborative learning support. *User Modeling and User-Adapted Interaction*, *19*(5), 387–431. http://doi.org/10.1007/s11257-009-9069-1

Walraven, A., Brand-Gruwel, S., & Boshuizen, H. P. A. (2008). Information-problem solving: A review of problems students encounter and instructional solutions. *Computers in Human Behavior*, *24*(3), 623–648.

Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: the role of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, *78*(3), 546–558.

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, *13*(1), 21–39.

Webb, N. M. (1991). Task-related verbal interaction and mathematics learning in small groups. *Journal for Research in Mathematics Education*, *22*(5), 366–389. http://doi.org/10.2307/749186

Webster, D. M., & Kruglanski, A. W. (1997). Cognitive and Social Consequences of the Need for Cognitive Closure. *European Review of Social Psychology*, *8*(1), 133–173. http://doi.org/10.1080/14792779643000100

West, T. V., & Kenny, D. A. (2011). The truth and bias model of judgment. *Psychological Review*, *118*(2), 357–378. http://doi.org/10.1037/a0022936

Winne, P. H., Hadwin, A. F., & Perry, N. E. (2013). Metacognition and computer-supported collaborative learning. In C. E. Hmelo-Silver, C. A. Chinn, C. A. A. Chan, & A. O'Donnell (Eds.), *International handbook of collaborative learning* (pp. 462–479). New York: Routledge.

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). San Diego, CA, US: Academic Press. http://doi.org/10.1016/B978-012109890-2/50045-7

Wopereis, I., Brand-Gruwel, S., & Vermetten, Y. (2008). The effect of embedded instruction on solving information problems. *Computers in Human Behavior*, *24*(3), 738–752.

Zelazo, P. D., & Cunningham, W. A. (2007). Executive function: Mechanisms underlying emotion regulation. In J. J. Gross (Ed.), *Handbook of emotion regulation* (pp. 135–158). New York, NY: Guilford Press.

Appendices

Appendix A - Translation of the six items capturing participants' intentions and expectations

Participants' intentions and expectations for a potential collaborative learning phase, including title and instructions on the screen. On the left side one can find the original German version, while on the right side the English translation is provided.

Handlungen und Erwartungen	**Intentions and Expectations**
Bitte lesen Sie jede der folgenden Aussagen durch und entscheiden Sie, wie stark diese Sätze *Ihre persönliche* Handlungen und Erwartungen in einer potentiellen gemeinsamen Lernphase ausdrücken.	Please read each of the following statements and decide how much they correspond to your personal intentions and expectations for a potential collaborative learning phase.
Bezugnehmend auf den Text mit dem Titel „Der Magen“ (*Beispiel*)	According to the text entitled „The stomach“ (*example*)
1. Wie viele Informationen über den Text werden Sie von ihrem Lernpartner erlangen?	1. How much information about the text will you receive from your learning partner?
Gar keine ; Sehr viele	None at all; Very much
2. Für wie wahrscheinlich halten Sie es, dass Ihr Lernpartner Ihnen die richtige Antwortmöglichkeit mitteilen wird?	2. In your opinion, how likely will your learning partner provide the correct answer item?
Sehr wahrscheinlich; Sehr unwahrscheinlich	Very likely; Very unlikely
3. Wie viele Informationen über den Text werden Sie Ihrem Lernpartner mitteilen?	3. How much information about the text will you provide to your learning partner?
Gar keine; Sehr viele	None at all; Very much
4. Wie intensiv werden Sie versuchen Ihr Verständnis vom Text zu begründen?	4. How strongly will you try to justify your understanding of the text?
Gar nicht; sehr intensiv	Not at all; Very strongly
5. Für wie wahrscheinlich halten Sie es, dass Sie Ihrem Lernpartner die richtige Antwortmöglichkeit mitteilen werden?	5. In your opinion, how likely will you provide the correct answer item to your learning partner?
Sehr wahrscheinlich; sehr unwahrscheinlich	Very likely; very unlikely
6. Wie intensiv werden Sie Ihren Lernpartner zu diesem Text befragen?	6. How deeply will you query your learning partner concerning the text?
Gar nicht; sehr intensiv	Not at all; very deeply

Appendix B – Initial article

Self-Directed Learning: A Cognitive and Computational Perspective

Todd M. Gureckis and Douglas B. Markant

A widely advocated idea in education is that people learn better when the flow of experience is under their control (i.e., learning is self-directed). However, the reasons why volitional control might result in superior acquisition and the limits to such advantages remain poorly understood. In this article, we review the issue from both a cognitive and computational perspective. On the cognitive side, self-directed learning allows individuals to focus effort on useful information they do not yet possess, can expose information that is inaccessible via passive observation, and may enhance the encoding and retention of materials. On the computational side, the development of efficient "active learning" algorithms that can select their own training data is an emerging research topic in machine learning. This review argues that recent advances in these related fields may offer a fresh theoretical perspective on how people gather information to support their own learning.

Appendix C - Encyclopedic texts about self-directed learning and machine learning

Self-directed Learning

Most adults spend a considerable time acquiring information and learning new skills. The rapidity of change, the continuous creation of new knowledge, and an ever-widening access to information make such acquisitions necessary. Much of this learning takes place at the learner's initiative, even if available through formal settings. A common label given to such activity is self-directed learning. In essence, self-directed learning is seen as any study form in which individuals have primary responsibility for planning, implementing, and even evaluating the effort. Most people, when asked, will proclaim a preference for assuming such responsibility whenever possible. (*The International Encyclopedia of Education (second edition), Oxford*)

Machine Learning/ Active Learning

Machine learning is a type of artificial intelligence that provides computers with the ability to learn without being explicitly programmed. Machine learning focuses on the development of computer programs that can teach themselves to grow and change when exposed to new data. These programs detect patterns in data and adjust program actions accordingly.

Active learning is a special case of semi-supervised machine learning in which a learning algorithm is able to interactively query the user (or some other information source) to obtain the desired outputs at new data points. In statistics literature it is sometimes also called optimal experimental design. (*Wikipedia*)

Appendix D - List of 60 scientific journal names for psychology/ learning sciences and computer/ cognitive sciences

List of journals, from which articles until 2013 were used as the scientific data base

Psychology/ Learning Sciences	Computer Sciences/ Cognitive Sciences
Psychological Science	Communications of the ACM
American Psychologist	Theoretical Computer Science
Annual Review of Psychology	Neural Computation
Psychonomic Bulletin & Review	Journal of Machine Learning Research
Organizational Behavior and Human Performance	Neurocomputing
Perception & Psychophysics	Neural Computing & Applications
Journal of Personality and Social Psychology	Computer Science & Engineering
Current Directions in Psychological Science	Bioinformatics
Developmental Psychology	Expert Systems with Applications
Psychological Reports	Computers & Education
Ecological Psychology	Educational Technology Research & Development
Educational Psychology Review	Journal of Educational Technology & Society
Frontiers in Psychology	Information Technology
Perspectives On Psychological Science	Journal of Information Systems Education
Journal of Applied Developmental Psychology	Journal of Machine Learning Research
Psychological Review	Computer Methods & Programs in Biomedicine
Journal of experimental Psychology: Applied	Communications

Journal of experimental Psychology: General	International Journal on Artificial Intelligence Tools
Journal of Mathematical Psychology	Engineering Applications of Artificial Intelligenc
Journal of experimental Social Psychology	Journal of Emerging Technologies in Web Intelligence
Cognition	NeuroImage
Cognition & Instruction	BMC Neuroscience
Journal of Cognition and Development	Neurocomputing
Cognitive Psychology	Neural Networks
Journal of experimental Psychology: Learning, Memory and Cognition	Neural Computing & Applications
Memory & Cognition	Neural Computation
Metacognition & Learning	Pattern Recognition
Topics in cognitive science	PLoS ONE
Trends in Cognitive Science	Neuron

Summary

Throughout all educational institutions and faculties students engage in different forms of cooperative learning. Whether in short-term knowledge-exchange-and-explanation-sessions or long-term problem-solving scenarios, learners need to cooperate in their educational life more than ever before. For a successful cooperation, they need to get aware of what and how good they and their learning partners understood the given learning material at hand. Only with an accurate awareness of the knowledge constellation within the group, learners might be able to provide or elicit helpful information for the present task. While scholars and teachers need a theoretical basis, which allows them to derive instructional or technological support to increase efficiency in cooperative learning, no models—especially no cognitive models—exist so far. Hence, the current dissertation likes to cover that research gap by providing a cognitive model from an individual perspective describing relevant processes of representing, monitoring and controlling own and others' knowledge states during cooperative learning situations.

Therefore, in a first step the dissertation develops a model termed MEDIA ("Metacognition in dyadic interaction"), which focuses on dyadic learning situations as the most basic form of group learning. The model argues that own *and* others' knowledge states can be represented cognitively and monitored by a process of evaluating current knowledge states with desired goal states. This means that learners in dyadic learning situations evaluate not only individual discrepancies between current states and goals states of knowledge (whether for their partner or for themselves), but also intragroup differences (the difference between own and partner knowledge level). According to the result of the evaluation, the model assumes that the learner will control for these differences by acts of providing or eliciting explanations. The model focuses therefore on processes of comprehension, comprehension evaluation as acts of monitoring, and on gaining comprehension through explanations as acts of controlling. Furthermore, the model claims that monitoring own and partner knowledge accurately should have an influence on the accuracy of providing and eliciting explanations (i.e. efficiency in cooperative learning), which will in turn increase learning effects in the dyad. Finally, processes of provided and received explanations should also have an influence on self- and partner-monitoring accuracy, as learners continuously have more cues available to evaluate own and partner knowledge states.

In a second step, the current dissertation reports on a set of three empirical studies testing some of the assumptions of the model presented above. The first two studies used a knowledge-exchange-and-explanation-setting as the cooperative learning scenario with low or high self-judged knowledge (i.e. monitoring) in combination with either low or high partner knowledge to test learners' control behavior in different knowledge constellations (i.e. providing or receiving explanations). First, results indicate that participants intend to provide explanations mainly based on the amount of own knowledge, while their expectations to receive explanations depend solely on the amount of partner knowledge. Second, the results showed that high-knowledgeable learners' discrimination abilities to distinguish between well and poorly understood learning items (i.e. relative self-monitoring accuracy) predicts efficient help-giving (i.e. relative control accuracy). But surprisingly, their low-knowledgeable partners could not benefit from this help-giving efficiency. Instead, low-knowledgeable partners mainly gained a better understanding through the amount of received explanations, and not through the efficiency of received explanations. Third, all learners showed low self-monitoring accuracy, but high partner-monitoring accuracy after the cooperation. A relation could be found between the amount of received explanations and learners' partner-monitoring accuracy, which partly explains the effect. Finally, learners were strongly influenced by their own knowledge representations when predicting partner knowledge, which is known as the false consensus bias. Testing the model in a more applied problem-solving-scenario, Study 3 investigated dyads holding either identical or complementary knowledge prior to an authentic, cooperative search task. Focusing on the effects of self- and partner-monitoring accuracy after the cooperation, results showed that dyads with complementary knowledge were more accurate in monitoring own knowledge, while no differences occurred between dyads of complementary and identical knowledge with regard to partner-monitoring accuracy. This time, self- and partner-monitoring accuracy was neither influenced by own nor by partner interaction behavior. However, Study 3 could confirm the false consensus effect found in Study 2.

Taken together, the current dissertation contributes to the need to better understand cooperative learning situations and how learners represent, monitor and control the given knowledge constellations within the group. As seen, the relationships between accurate monitoring, accurate control behavior and learning gains are not trivial. Moreover, cognitive processes seem to be biased by a belief in a false knowledge-consensus within the dyad. With this dissertation, a first promising step was made towards a cognitive view on regulation processes in cooperative learning, which needs to be continued in further researcher endeavors.

Zusammenfassung

Über alle Bildungsinstitutionen hinweg und fächerübergreifend ist eine didaktische Form der Wissensvermittlung kaum noch wegzudenken: kooperatives Lernen. Ob in kurzfristigen Lernsessions (z.B. kurzes Selbststudium mit anschließenden gegenseitigen Erklärungen) oder langfristigen Problemlösungsszenarien, die Lernenden müssen im Laufe ihres Bildungswegs mehr denn je miteinander kooperieren. Für eine erfolgreiche Zusammenarbeit ist das Bewusstsein um das was und wie gut Lerner und ihre Lernpartner das vorgegebene Lernmaterial verstanden haben von immenser Wichtigkeit. Nur mit einem genauen Bewusstsein um die Wissenskonstellation innerhalb der Gruppe sind die Lernenden in der Lage, passende Informationen bereitzustellen oder fehlende Informationen zu erfragen. Um kooperatives Lernen effizient zu gestalten, benötigen Wissenschaftler und Lehrer eine theoretische Grundlage, die es ihnen ermöglicht, den Lernenden didaktisch oder technologisch sinnvolle Lernumgebungen bereitzustellen. Eine solche theoretische, vor allem kognitive Grundlage ist bisher nicht bekannt. Aus diesem Grund möchte die Dissertation diese Forschungslücke schließen und ein Modell liefern, welches die relevanten kognitiven Prozesse der Repräsentation, Überwachung und Kontrolle von eigenen und anderen Wissensständen erklärt.

In einem ersten Schritt wurde dazu ein Modell entwickelt, das die individuell-kognitiven Prozesse eines Lerners in der grundlegendsten Form des Gruppenlernens, in dyadische Lernsituationen adressiert – genannt "Metakognition in dyadischen Interaktionen". Das Modell postuliert, dass die Wissensstände des Lerners *und* des Lernpartners zunächst repräsentiert und durch einen Prozess der Bewertung von aktuellen Wissenszuständen mit den gewünschten Zielzuständen überwacht werden können. Dies bedeutet, dass die Lernenden in dyadischen Lernsituation nicht nur individuelle Diskrepanzen zwischen den aktuellen Zuständen und gewünschten Zielen (ob für ihren Partner oder für sich selbst), sondern auch intragruppale Unterschiede evaluieren können – also den Unterschied zwischen dem eigenen und dem Partnerwissen. Im Anschluss an die Evaluation geht das Modell davon aus, dass der Lernende diese Unterschiede durch Bereitstellung oder Erfragung von Erklärungen kontrollieren kann. Das Modell konzentriert sich daher auf die Verständnis- und Verständnisbewertung als Handlungen der Überwachung und auf die Verständnisgewinnung durch Erklärungen als Kontrollhandlungen. In Anlehnung an die Metakognition in Individuen wird vermutet, dass je akkurater die Überwachung des eigenen Wissens und des Partnerwissens ist, desto größer die Genauigkeit nur für die Lernelemente Erklärungen bereitzustellen, für die der Partner diese benötigt, und nur für die Lernelemente Erklärungen zu erfragen, für die der Lerner auch wirklich welche benötigt. Steigt mit der Akkuratheit der Überwachung auch die Akkuratheit der Kontrollhandlungen, sollte sich dies auf die Lernwirksamkeit in der Dyade auswirken. Abschließend beschreibt das Modell, dass die Prozesse des Gebens und Nehmens von Erklärungen ebenfalls einen Einfluss auf die Genauigkeit der Selbst- und Partnerüberwachung haben sollte, da die Lernenden kontinuierlich mehr Hinweise zur Beurteilung der Wissensstände erhalten.

In einem zweiten Schritt berichtet die vorliegende Dissertation über ein Studienprogramm von drei empirischen Studien, die durchgeführt wurden um einige der oben dargelegten Annahmen des Modells zu untersuchen. Die ersten beiden Studien nutzten dabei ein klassisches Lernsetting mit Selbststudium und gegenseitigen Erklärungen als kooperatives Lernszenario. Manipuliert wurden die Selbstbeurteilung des Wissens (mit niedrigen oder hohen Selbstbeurteilungen) in Kombination mit niedrigem oder hohem Partnerwissen, um das Kontrollverhalten der Lernenden in unterschiedlichen Gruppenkonstellationen zu testen. Es zeigte sich erstens, dass die Teilnehmer Erklärungen vor allem auf der Grundlage der Menge des eigenen Wissens anbieten, während ihre Erwartungen Erklärungen zu erhalten allein von der Höhe des Partnerwissens abhing. Zweitens zeigte sich, dass Lernende mit einer hohen Selbstbeurteilung eine bessere Fähigkeit besaßen, gut- und schlecht-verstandene Lernelemente (d.h. relative Selbstüberwachungsgenauigkeit) zu unterscheiden. Dies führte zu einer effizienteren Hilfestellung für Partner mit niedrigen Selbstbeurteilungen (d.h. eine relative Kontrollgenauigkeit), von der sie jedoch überraschenderweise wenig profitieren konnten. Im Gegensatz zu den vermuteten Wirkzusammenhängen haben Lernpartner mit niedrigen Selbstbeurteilungen vor allem durch die Menge der erhaltenen Erklärungen und nicht durch die Effizienz der erhaltenen Erklärungen ihre Lernleistungen steigern können. Drittens zeigten alle Lernenden eine geringe Selbstüberwachungsgenauigkeit, aber eine hohe Partnerüberwachungsgenauigkeit nach der Zusammenarbeit. Während Lernenden durch ihre gegebenen Erklärungen nicht zu einer akkurateren Selbstbeurteilung gelangten, scheinen erhaltene Erklärungen dazu beigetragen zu haben, dass die Lernenden eine genauere Repräsentation des Partnerwissens erlangten. Schließlich zeigten die Studien, dass Lernende stark von ihren eigenen Wissensrepräsentationen beeinflusst sind, wenn sie das Partnerwissen beurteilen sollen, was in der Social-Cognition-Literatur auch als falscher Konsens bekannt ist. Bei der Untersuchung des Modells in einem mehr angewandten Problemlösungsszenario untersuchte Studie 3 Dyaden mit identischem oder komplementärem Wissen, die eine authentische, kooperative Suchaufgabe lösen mussten. Der Fokus in dieser Studie lag auf der Akkuratheit der Selbst- und Partnerüberwachung nach der Kooperation. Die Ergebnisse zeigten, dass Dyaden mit komplementärem Wissen akkurater in der Überwachung des eigenen Wissens waren, während keine Unterschiede zwischen Dyaden komplementärem und identischem Wissens in Bezug auf die Akkuratheit der Partnerüberwachung auftraten. Im Gegensatz zu Studie 2 zeigte sich kein Einfluss des Lerner- oder Partnerverhaltens auf die Selbst- und Partnerüberwachung. Allerdings konnte Studie 3 den Bias des falschen Konsens bestätigen.

Zusammengenommen trägt die vorliegende Dissertation dazu bei, die kooperativen Lernsituationen besser zu verstehen und die kognitiven Zusammenhänge von Repräsentation, Überwachung und Kontrolle eigener wie fremder Wissensstände darzulegen. Es zeigt sich, dass die Zusammenhänge zwischen akkurater Lernüberwachung, effizientem Kontrollverhalten und Lernerfolg nicht trivial sind und durch einen egozentrischen Effekt verzerrt werden. Mit dieser Dissertation wurde ein erster vielversprechender Schritt zur kognitiven Betrachtung von Regulierungsprozessen im kooperativen Lerngruppen gemacht, die durch weitere Forschungsarbeiten fortzusetzen ist.

Danksagung

Die Art und Weise wie Menschen miteinander kommunizieren, sich austauschen und voneinander lernen, hat von je her eine große Faszination auf mich ausgeübt und diese Dissertation maßgeblich beeinflusst und geleitet. Den Anderen zu verstehen, dessen Verständnis mit dem eigenen Verständnis abzugleichen, Konsequenzen zu erörtern und, wo möglich, eine Verständnisebene zu finden, sind Fähigkeiten, die ich, über die in dieser Dissertation eng gefassten Lernsituation hinaus, mehr denn je sehr zu schätzen weiß. Mein Dank gilt daher an erster Stelle Jürgen Buder, der durch unzählige Gespräche, viele Ideen, immer neue Fragen und einer nicht näher zu bestimmenden Anzahl an gemeinsamen Zigaretten-Denk-Pausen maßgeblich zu einer gemeinsamen, aber vor allem allgemeinen Verständnisebene dieser Dissertation beigetragen hat.

Des Weiteren gilt mein Dank Prof. Hesse und Prof. Bodemer sowie allen ehemaligen und heutigen Mitarbeitern ihrer Arbeitsgruppen; in ihnen fand ich einen fruchtbaren wissenschaftlichen Nährboden und offenen Raum des Diskurses, der viele wertvolle Anregungen hervorgebracht hat.

Abseits dieser hoch rationalen Objekt- und Metaebene bin ich allen Menschen aus tiefsten Herzen dankbar, die meine Promotionszeit durch sportliche Aktivitäten, kulinarische Genüsse, kreative Spinnereien, ausgelassenen Musikabende und vielen Umarmungen sehr bereichert haben. Ihr habt Tübingen und Stuttgart zu meinem Zuhause gemacht!

Ein besonderer Dank gilt meinen Eltern, auf deren Unterstützung ich jederzeit zählen konnte. Diese Sicherheit ist allzu oft unterschätzt. Und zu guter Letzt möchte ich mich bei Jana bedanken, die mich gelehrt hat, dass Menschsein sicherlich nicht eine Verständnissuche, keine Reflexion und ganz sicher nicht in irgendeinen Satz zusammengefasst und (hier) beschrieben werden kann.