UNIVERSITÄT DES SAARLANDES

Fakultät HW

Bereich Empirische Humanwissenschaften

Co-Present Learning With Tablets in Primary School

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Philosophie
der Fakultät HW
Bereich Empirische Humanwissenschaften
der Universität des Saarlandes

vorgelegt von

Lara Johanna Kataja

aus Saarbrücken

Bonn, 2021

Dekan: Prof. Dr. Jörn Sparfeldt

Berichterstatter:

Professor Dr. Armin Weinberger, Universität des Saarlandes

Pantelis Papadopoulos, Associate Professor, PhD, University of Twente

Kristine Lund, HDR, PhD, Ecole Normale Supérieure de Lyon

Tag der Disputation: 22.01.2021

Acknowledgements

First of all, I would like to thank my supervisor and co-author Professor Armin Weinberger. Thank you for giving me the opportunity to do doctoral studies and to be a member of this department. Thank you for accompanying and guiding me along this path, for providing the time and support it takes to navigate one's way through the world of research, to write articles, chapters, and conference contributions. Thank you for discussing my work with me, and for giving me feedback and directions.

Further, I want to thank Dr. Jochen Rick and Dr. Dimitra Tsovaltzi. Jochen, as the creator of the Proportion app, most of this research would not have been possible without you. I am also grateful for all your remote support from USA and that we managed to successfully publish together. Dimitra, you know my highs and lows during this PhD student time very well. I am so glad to do research with you and write an article together, as well as for our good discussions.

To my other present and past colleagues, Thomas Puhl, Allison Kolling, Yvonne Berkle, and Raluca Judele: Thank you all for our interesting conversations, fun time, and mutual support. I think you all make a great team and an awesome work environment. Professor Pantelis Papadopoulos, you visited our department several times, and in each of these visits, you took time to discuss with me and provide very valuable advice, while being a very pleasant and fun visitor. Thank you so much. I would also like to thank the numerous student assistants who were involved in the research during my PhD student time at one point or another: Angeliki Ziaka, Bianca Wiehr, Christian Backes, Cindy Pérez, Daniel Kopp, Diana Salazar, Eirini Papanastasiou, Georgia Patinioti, Jan Lietz, Jil Thurmes, Johanna Gröning, José Cuida, Luka Philippi, Miguel Rejón, Tim Steuer, and Yvonne Berkle. Thank you for your great support, and for the positive experience to work, collaborate, and learn with you.

I would also like to thank the cooperating primary schools in Saarland, the teachers, children and their parents, for the opportunity of collecting the data for this thesis. Moreover, I am grateful towards the German Academic Exchange Service (DAAD) for funding my attendance at the CSCL conferences 2017 and 2019.

Last but not least, to my family and friends: In particular, I want to thank Nadja Simon and Joonas Kataja for helping me spell-check the final version of this thesis. Dear family, thank you for everything; thank you for listening, for your encouragement, for your imperturbable trust in me. To Joonas, kiitos paljon, I am so happy to have you.

Table of Contents

Inhalt

List of Figures

List of Tables

List of Abbreviations

ANOVA	Analysis of variance (statistics)
App	Software application, for example on a tablet
BiCo	Bipolar continuous rating scale (Schmitt, Rick, & Weinberger, 2019)
CI	Confidence interval (statistics)
CSCL	Computer-supported collaborative learning
CV	Coefficient of variation (statistics; Weinberger, Stegmann, & Fischer, 2007)
d	Cohen's d (effect size, statistics)
F	F-value (statistics)
M	Mean (average; statistics)
MANOVA	Multivariate analysis of variance (statistics)
MOOC	Massive open online course
n	Sample size (statistics)
p	p-value / probability (statistics)
PSS	Perceptual symbol systems (Barsalou, 1999)
r	Pearson correlation coefficient (statistics)
SD	Standard deviation (statistics)
STEP	Science through Technology Enhanced Play (Danish, Enyedy, Saleh, Lee, & Andrade, 2015)
STRAT	Strategy prompts (Schmitt & Weinberger, 2019)
t	t-value (statistics)
VERB	Verbalization prompts (Schmitt & Weinberger, 2019)
α	Krippendorff's alpha (statistics; Krippendorff, 2012)
η^2	Eta-squared (effect size; statistics)

Overall Structure and List of Publications

The present doctoral thesis on children's co-present learning with tablets in primary school is based on four publications, and comprises eight chapters in total: Chapters 1 – 4 elaborate on the general theoretical background of this doctoral research program; chapters 5 and 6 discuss methodological considerations and the learning environment of this research program; chapter 7 presents the four studies of this research program in summarized form; and chapter 8 provides a general discussion and outlook based on this research program. Three of the studies presented in chapter 7 are published, either as chapter in an international peer-reviewed handbook (study I), or as articles in international peer-reviewed journals (studies II and III); study IV presented in chapter 7 is currently in preparation.

Study I

Schmitt, L. J., & Weinberger, A. (2018). Computer-supported collaborative learning: Mediated and co-present forms of learning together. In J. Voogt, G. Knezek, R. Christensen, & K.-W. Lai (Eds.), *Second handbook of information technology in primary and secondary education* (pp. 217-231). Cham, Switzerland: Springer. doi:10.1007/978-3-319-71054-9_15

Study II

Schmitt, L. J., Rick, J., & Weinberger, A. (2019). BiCo: A bipolar continuous rating scale for children's technology evaluation. *Technology, Pedagogy and Education, 28*(5), 503–516. doi:10.1080/1475939X.2019.1661279

Study III

Schmitt, L. J., & Weinberger, A. (2019). Fourth graders' dyadic learning on multi-touch interfaces - versatile effects of verbalization prompts. *Educational Technology Research and Development, 67*(3), 519-539. doi:10.1007/s11423-018-9619-5

Study IV

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

In addition to the publications listed above, I presented the research done in the course of the present thesis at multiple conferences:

Rick, J., Kopp, D., Schmitt, L., & Weinberger, A. (2015). Tarzan and Jane share an iPad. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 1* (pp. 356–363). Gothenburg, Sweden: The International Society of the Learning Sciences. Abstract retrieved from https://www.isls.org/cscl2015/

Schmitt, L., Rick, J., & Weinberger, A. (2016, March). *BiCo: Eine bipolare kontinuierliche Ratingskala zur Minderung von Deckeneffekten bei der Evaluation von Bildungstechnologien durch Kinder*. Paper presented at the GEBF conference, Berlin, Germany.

Schmitt, L. J., Rick, J., & Weinberger, A. (2017, August / September). *Capturing children's evaluation of educational technologies with "BiCo"*. Paper presented at the EARLI conference, Tampere, Finland. Abstract retrieved from https://www.earli.org/earli-2017

Schmitt, L., Tsovaltzi, D., & Weinberger, A. (2019). Effects of process heterogeneity in collaborative embodied learning with tablets. In D. Tsovaltzi & A. Weinberger (Chair and Co-Chair), *Group formation in the digital age: Relevant characteristics, their diagnosis, and combination for productive collaboration*. Symposium conducted at the CSCL conference. In K. Lund, G. P. Niccolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019, Volume 2* (pp. 719-726). Lyon, France: The International Society of the Learning Sciences. Abstract retrieved from https://cscl2019.com/

Schmitt, L., & Weinberger, A. (2016, August). *Dyadic argumentation of elementary school children with a reflective tool*. Paper presented at the EARLI SIG 20 & 26 meeting, Ghent, Belgium.

Schmitt, L. J., & Weinberger, A. (2017). Collaborative learning on multi-touch interfaces: Scaffolding elementary school students. In B. K. Smith, M. Borge, E. Mercier, & K. Y. Lim (Eds.), *Making a Difference: Prioritizing Equity and Access in CSCL, 12th*

International Conference on Computer Supported Collaborative Learning (CSCL) 2017, Volume 1 (pp. 9-16). Philadelphia, PA: The International Society of the Learning Sciences. Abstract retrieved from https://cscl17.wordpress.com/

Schmitt, L. J., & Weinberger, A. (2017, August / September). *Scaffolding elementary school students' collaborative learning on tablets*. Paper presented at the EARLI conference, Tampere, Finland. Abstract retrieved from https://www.earli.org/earli-2017

Weinberger, A., & Schmitt, L. (2016, March). *Produktive Divergenz beim kooperativen Lernen mit iPads – eine multimodale Analyse des Falls "Tarzan und Jane"*. Paper presented at the GEBF conference, Berlin, Germany.

Abstract

The present thesis investigates co-present collaborative and embodied learning with a touch interface (tablet) in a primary school context, in the domain of mathematics learning (proportional thinking). To this end, this thesis elaborates on theoretical and methodological considerations relevant to this research topic, namely computer-supported collaborative learning (CSCL), embodied cognition and embodied learning environments, multifaceted learning processes and their significance for learning, the domain of proportional reasoning and the *Proportion* learning environment, and last, research methods for capturing children's experiences and learning processes in this specific context. The four studies presented in this thesis provide different angles on the main topic of co-present learning with tablets: Study I (Schmitt & Weinberger, 2018) focuses on computer-supported collaborative learning research and learning scenarios in primary and secondary education, both computer-mediated and co-present, and identifies challenges, such as coordination issues or overly playful behavior, along with strategies to address these challenges, such as scaffolding learners. Study II (Schmitt, Rick, & Weinberger, 2019) focuses on the development of methods for research with children; based on the findings that currently available survey tools often do not suffice in adequately capturing children's experience of a technology intervention, a new tool (*BiCo*) is being developed and tested, with the obtained results being in support of this new instrument. Study III (Schmitt & Weinberger, 2019) investigates the effects of strategy and verbalization prompts in collaborative embodied learning, with the key finding that verbalization prompts have diverse effects, like enhancing the quality of dialogue, but also increasing the frequency of off-task behavior and negative emotions. Further, study IV (Schmitt, Tsovaltzi, & Weinberger, in preparation) explores the role of learning process heterogeneity on a cognitive and bodily dimension, suggesting that both homogeneous as well as heterogeneous learning processes can be advantageous depending on the dimension of interest. For example, while heterogeneous bodily processes improved performance in the app, homogeneous cognitive processes fostered knowledge convergence within the group. Upon presenting these four studies, the results are generally discussed, and directions for future research are provided, such as looking into the possible role of bodily transactivity, or developing methods for automatic analysis of learning processes.

Keywords: co-present learning, embodied cognition, touch screen interfaces, rating scales, scaffolding, heterogeneity, learning processes

Zusammenfassung

Die vorliegende Doktorarbeit untersucht kopräsentes kollaboratives und verkörperlichtes Lernen mit einem Touchscreen-Interface (Tablet) im Grundschulkontext im Bereich des Mathematiklernens (proportionales Denken). Dazu elaboriert die vorliegende Arbeit theoretische und methodische Überlegungen, die relevant für dieses Forschungsthema sind, nämlich computer-unterstütztes kollaboratives Lernen (computer-supported collaborative learning, CSCL), verkörperlichte (embodied) Kognition und körperliche Lernumgebungen, verschiedene Facetten von Lernprozessen und deren Bedeutung für das Lernen, die Domäne des proportionalen Denkens und die *Proportion*-Lernumgebung, und schließlich Forschungsmethoden zur Erfassung der Erfahrungen der Kinder und der Lernprozesse in diesem spezifischen Kontext. Die vier in dieser Arbeit präsentierten Studien bieten unterschiedliche Zugänge zu diesem Hauptthema des kopräsenten Lernens mit Tablets: Studie I (Schmitt & Weinberger, 2018) fokussiert sich auf Forschung und Szenarien des computer-unterstützten kollaborativen Lernens im Primar- und Sekundarbereich, sowohl computer-mediiert als auch kopräsent, und identifiziert Herausforderungen, wie Koordinationsbedarf oder übermäßig spielerisches Verhalten, sowie Strategien diesen Herausforderungen zu begegnen, wie zum Beispiel Scaffolding der Lernenden. Studie II (Schmitt, Rick, & Weinberger, im Druck) fokussiert sich auf die Entwicklung von Methoden zur Forschung mit Kindern; basierend auf Befunden, dass die gegenwärtig verfügbaren Fragebogen-Instrumente bezüglich der adäquaten Erfassung der Erfahrungen von Kindern bzgl. einer Technologieintervention nicht genügen, wird ein neues Instrument (*BiCo*) entwickelt und getestet, wobei die gewonnenen Ergebnisse dieses neue Instrument stützen. Studie III (Schmitt & Weinberger, 2019) untersucht die Effekte von Strategie- und Verbalisierungsprompts im Zusammenhang mit kollaborativem verkörperlichtem Lernen, mit den zentralen Befunden, dass Verbalisierungsprompts verschiedenartige Effekte haben, wie die Verbesserung der Dialogqualität, aber auch ein erhöhtes Vorkommen von Off-Task-Verhalten und negativen Emotionen. Weitergehend exploriert Studie IV (Schmitt, Tsovaltzi, & Weinberger, Manuskript in Vorbereitung) die Rolle von Lernprozessheterogenität auf einer kognitiven und einer körperlichen Dimension, und legt nahe, dass sowohl homogene als auch heterogene Lernprozesse vorteilhaft sein können, je nachdem welche Dimension von Interesse ist; zum Beispiel verbesserten heterogene körperliche Prozesse die Leistung in der App, während homogene kognitive Prozesse Wissenskonvergenz innerhalb der Gruppe förderten. Anschließend an die Präsentation dieser vier Studien werden die Ergebnisse allgemein diskutiert und Denkrichtungen für zukünftige Forschung dargelegt, wie zum Beispiel die

Betrachtung der potentiellen Rolle von körperlicher Transaktivität, oder die Entwicklung von Methoden der automatischen Analyse von Lernprozessen.

Keywords: ko-präsentes Lernen, verkörperlichte Kognition, Touchscreen-Interfaces, Beurteilungsskalen, Scaffolding, Heterogenität, Lernprozesse

1. Introduction – Co-Present Learning With Tablets in Primary School

Nowadays, digital technologies are present everywhere, and researchers are exploring how technologies impact our work life, our free time, our health, and also what role they may have for teaching and learning. Educational Technology research aims at expanding our knowledge about how people learn: What role do different technologies have in the learning process and how do they shape learner interactions or individual processes? How do technologies support or hinder communication, collaboration, or knowledge acquisition? To understand and enhance learning with technologies, it is essential to build and refine theories and best practices of computer-supported learning, which then can serve as a basis for developing sensible new technologies for learning.

A multitude of different technologies for supporting individual or collaborative learning activities are available for use in informal and formal learning settings, such as classrooms. However, considering the pervasiveness of technologies in our daily lives, technologies are utilized comparatively rarely in schools (Deutsche Telekom Stiftung, 2015). According to a recent survey, less than half of all teachers in Germany (47.6 %) utilize technology in their teaching at least once per week, and 7 % of all teachers never make use of any sorts of digital media in the classroom (Deutsche Telekom Stiftung, 2015). Doubts regarding the general benefits or risks of technologies for learning, or lack of experience regarding the concrete integration of technologies into a learning scenario may stop educators from using technologies in their classrooms. In addition to organizational factors, such as availability of devices, one crucial prerequisite for a successful integration of technologies into education and learning activities is to prepare teachers adequately and support them in changing their teaching practices towards productively utilizing technologies (Deutsche Telekom Stiftung, 2015). In Germany, it is currently planned to spend five billion Euro for digitalization in schools, with the goal of supporting sufficient technological equipment, as well as matching pedagogical concepts, and adequately trained teachers (Die Bundesregierung, 2019). With all the money being made available for technologies in schools and for teacher training, it is also fundamental to continuously build the scientific basis of theories on teaching and learning.

How do people learn? Educational materials traditionally included text books with symbolic representations (written words / letters, numbers and mathematical notations). When technologies supporting a range of different representation modes became more common, multimedia learning materials combined information on different modes, for instance, verbal information (e.g., audio or text) combined with non-verbal information (e.g., graphics or video) (Moreno & Mayer, 1999), under the assumption that information presented on different modes is processed in different processing channels, and that learning is more effective when both channels are utilized, like combining graphics (visual channel) with spoken text (verbal channel) (Mayer, 2008). This form of multimedia learning based on the dual-channel theory has been well researched with substantial support for the modality principle (Mayer, 2008). In the last years, researchers started to consider if the *body* or *bodily activities* such as gestures could serve as an additional channel or mode of representing information, reducing cognitive load (e.g., Cook, Mitchell, & Goldin-Meadow, 2008; Ottmar & Landy, 2017).

Collaborative learning – learning activities in small groups of peers – is a prominent paradigm in the learning sciences and often associated with learning success, given productive interactions within the groups, which can be facilitated through the instructions or the learning environment design (Jeong & Hmelo-Silver, 2016). *Computer-supported collaborative learning (CSCL)* regards collaborative learning that is facilitated or mediated by different sorts of technologies. CSCL has a long tradition in research on physically separated learners constructing knowledge online through argumentative discourse, usually via written texts in forums or social media (e.g., Schellens, Van Keer, De Wever, & Valcke, 2007; Tsovaltzi, Judele, Puhl, & Weinberger, 2015). In general, also outside of educational settings, a lot of our (digitally enriched) interaction and communication is nowadays mediated through screens, for example, through messaging services on smartphones. Despite the advantages of being able to communicate anywhere and anytime, one can also raise concerns regarding this remote and possibly artificial way of communication. These concerns may be one reason for motivating the development of interaction and learning scenarios which bring the participants together again physically. Indeed, beyond exchanging texts online, CSCL also regards co-present learning scenarios, with learners being physically present with each other, supported by technology (e.g., Mercier, Higgins & da Costa, 2014; Roschelle & Teasley, 1995; Schmitt & Weinberger, 2018 / study I, section 7.1). In co-present CSCL scenarios, the bodily aspects of learning and interacting come into focus, as the learners do not communicate in a mediated way (as in exchanging texts online), but are bodily co-present.

One way of supporting co-present CSCL are interactive devices with touch screens, for instance, tablets or smartphones. These sorts of devices have been very popular and in widespread use in the recent years, at least when it comes to private use. In 2018, 47.5 % of all households in Germany were equipped with at least one tablet computer, and 77.9 % with at least one smartphone (Statistisches Bundesamt, 2018). Interacting with touch screens appears to be highly intuitive and natural: Children as young as three years old (Hiniker et al., 2015), and even orangutans (e.g., Tabor, 2000) can successfully interact with them.

Researchers, educators and teachers, as well as the general public became interested in the effects of tablets on our (human) society in general, and learning and education specifically. For example, this interest is reflected in the big newspapers in Germany which were picking up the topic of tablets in school with headlines like “Das wischende Klassenzimmer” (translated: “The swiping classroom”; Himmer, 2015), or more critically “Zwischen Mobbing und Medienkompetenz” (translated: “Between bullying and media competency”; Haas, 2018), to mention just a few. Among the reported advantages that tablets in schools may have, are their capability to easily store information or homework, as well as opportunities for communication (Himmer, 2015), or increased interactivity and learners’ attention, as well as support of discussions and collaborative activities (“Das digitale Klassenzimmer,” 2018). However, concerns have also been raised, for instance, that tablets can be a source for unproductive distractions (“Das digitale Klassenzimmer,” 2018), or a tool for bullying classmates (Haas, 2018). Learning with tablets as a hot topic in the 2010s deserves to be looked at through a scientific lens in order to assess its potential merits and risks, and to advance theories of how learners interact with a touch device, and how these processes are shaped by different conditions like the sort of instructions or the way of grouping learners into a team.

One key feature of *multi-touch tablets*, compared to traditional computers with keyboard and mouse, is their natural and direct input via (multiple) fingers (Rick, 2012a). This implies that, beyond dealing with symbolic representations, the interaction with touch screens also allows for direct manipulation of digital objects (Rick, 2012b; Rick, Bejan, Roche, & Weinberger, 2012), which may be in support for the *embodiment* aspect of cognition and learning (Schneps et al., 2014). For example, using natural finger gestures for zooming in and out, learners can explore the dimensions of space on a tablet (Schneps et al., 2014).

What is embodiment? I would like to start with an example of an everyday experience (for more examples, see Barsalou, 1999, 2008): When following a personally meaningful sports event or competition in TV, for instance figure skating, the viewer may experience something that goes beyond merely watching the athlete’s movements on the screen. Rather than watching, the experience may be characterized as feeling, or *embodying* the movements. This experience may be even more an embodied one, if the viewer has a personal connection to the respective sports, for instance, practices it themselves actively (or used to do so). Based on theories of embodied cognition (see chapter 3), the viewer may have formed multimodal representations of the sports movements in the brain as a result of physically exercising the sports. And these representations may be accessed and simulated while watching a related event or competition on the TV screen. In terms of neural activity, action and mental simulation of action may not be very different (e.g., Barsalou, 2008; Martin, 2007).

Can embodiment be harnessed for learning? I would like to illustrate this with one more everyday life example: At some point during high school, we covered the topic of how the heart looks and functions. Our biology teacher then brought a real pig heart to class to make us ‘grasp’ the topic better. We were supposed to inspect the heart, hold it in our hands, touch and feel it, and look at it from all angles. Approximately 15 years later, I still have a vivid memory of this lesson and appreciate our teacher’s effort to make an abstract topic ‘graspable’ and to experience it with all senses.

What if this embodying of an activity or a concept (so deeply that you can still easily access it many years later) could be applied for acquiring knowledge and skills in – rather abstract – topics like mathematics? Could children equipped with interactive tablets, that afford to engage with them through bodily movements, similarly build robust multimodal representations? Could a bodily, immersive learning activity help them to grasp mathematical concepts, to feel them, to *embody* them?

The ministry of education in Saarland[1] explicitly acknowledges the benefits of getting children moving in the classroom and involving bodily experiences in learning activities, and supports the development of information material for teachers, for example, the brochure “Bewegte Schule” (translated “The moving school”) by Paulus, Gigout, Kaczmarek, and Wydra (2011). In this brochure, Paulus et al. (2011) write about positive effects of moving on cognitive and social learning, and on motivation, among other benefits of bodily activity, such as increased health and fitness. Paulus et al. (2011) further argue that children should get moving not only in the breaks or in dedicated sports classes, but also during classes in subjects like mathematics or German, so that bodily experiences may be utilized as another channel for information processing and learning, in addition to hearing and seeing.

Under the assumption that learning is a social and active process of constructing knowledge (rather than passive reception of information), and that movements and bodily experiences (rather than reading formal notations or symbols) may support learning, this thesis utilizes a learning environment (*Proportion* app; e.g., Rick, 2012a) which runs on a tablet computer and enables the users to actively manipulate representations of quantities on the screen through direct input via the body (the fingers). By bodily engaging with this learning environment,

[1] Saarland is the federal state in Germany in which this thesis has been developed.

young children (about 10 years of age) may acquire a better understanding of proportional relations, i.e., improve their proportional reasoning. Proportional reasoning is an important but difficult domain of mathematics (e.g., Boyer, Levine, & Huttenlocher, 2008), which however is not covered in depth in primary school curricula in Saarland (Saarland Ministerium für Bildung, Familie, Frauen und Kultur, 2009). *Proportion* allows learners to try out different inputs, to interact with the representation, and it provides fast and accurate feedback to learners. The *Proportion* learning scenario provides that two participants share one tablet, therefore, it enables collaborative interactions (Rick, 2012a). Bodily experiencing proportional relations coupled with co-present collaborative learning implies a fairly rich learning scenario: The learners do not only manipulate the objects on the screen and embody the proportional relations for themselves, but they can also perceive (and react to) the learning partner's movements or overtly expressed emotions (see also, Rick, 2012b). Two learners sharing one tablet may establish a joint focus regarding the shared representation, and guide the partner's attention by performing adequate movements or through pointing gestures. Moreover, learners may closely coordinate their bodily behavior to advance in the app, and they have the opportunity to verbally discuss and reflect together on these bodily experiences. While reflecting on learning material is also possible individually, "the presence of a partner provides a natural context for elaborating one's own reasoning" (Teasley, 1997; p. 379). From a researcher's perspective, investigating such a rich learning scenario also implies a need for a multimodal analysis of the manifold verbal and non-verbal learning processes; consequently, the present thesis regards the quality of learners' verbalizations, but also their bodily expressed task focus and emotional processes.

Reflecting on the development of the learning sciences as a whole, Duschl (2008) identified key constructs that are being worked on in order to transform learning environments: "(a) transition from novice to expert performance, (b) using prior knowledge, (c) scaffolding, (d) externalization and articulation, (e) reflection, and (f) building from concrete to abstract knowledge" (Duschl, 2008, p. 272). Especially the constructs (c), (d), (e), and (f) are central to the work done in the present thesis.

The overall scope of the present thesis is children's co-present collaborative learning with tablets as facilitators for a collaborative and embodied approach to knowledge construction. To this end, the four studies of this thesis will approach this topic from different angles. The first approach is a *conceptual* one, introducing the topic of CSCL research and scenarios, and specifically also discussing co-present collaborative learning scenarios for young learners (Schmitt & Weinberger, 2018 / study I, section 7.1). The second approach is a *methodological* one; rather than the children being merely passive objects of the research, research in the context of children learning with technologies needs also to include subjective experiences from the children themselves directly. Therefore, *BiCo* as a child-friendly survey method was developed and investigated in study II (Schmitt, Rick, & Weinberger, 2019; section 7.2). The third approach to this topic is an *experimental* one, with a focus on the conditions of successful learning in the given context and the refinement of theoretical considerations, regarding questions of supporting learners' collaborative and embodied learning with additional prompts (Schmitt & Weinberger, 2019 / study III, section 7.3), and the role of heterogeneous learning processes on multiple dimensions in collaborative groups (Schmitt, Tsovaltzi, & Weinberger, in preparation / study IV, section 7.4).

Summarized, in the four studies done as part of this thesis, the following guiding research questions will be addressed:

- **Study I**: What are typical scenarios of children's computer-supported collaborative learning, both computer-mediated and co-present? What potential challenges are associated with different forms of scenarios, and how can these challenges be addressed to support productive learning processes?
- **Study II**: What are the strengths and weaknesses of currently available survey instruments to measure children's evaluation of an educational technology? How could an appropriate instrument be designed, and to what extent is the newly developed BiCo instrument a useful tool for quantitatively assessing children's evaluations?
- **Study III**: How do additional verbalization and strategy prompts impact multifaceted learning processes as well as learning outcomes in collaborative embodied learning with the *Proportion* tablet application? What are the effects of combining bodily activities with verbal reflection and prompts for strategic behavior?
- **Study IV**: To what extent does within-dyad heterogeneity in learning processes further shape learning in collaborative embodied learning scenarios? What is the role of bodily and cognitive process heterogeneity with respect to performance and cognitive processes variables?

Rather than contrasting basic features of a learning scenario, such as collaborative vs. individual learning, or bodily activities vs. no bodily activities, the present thesis aims at identifying the nature and conditions of learning processes within a given scenario. Studies III and IV of this thesis are conducted within the paradigm of collaborative embodied learning. The goal is to put forward theories of *how* learning processes operate, *how* they relate to each other, and *how* they can be supported. While in these studies, the factors of embodiment and collaborative learning are not being experimentally varied, both aspects define the context in which this research takes place. Therefore, to ground the studies into this broader research context, the underlying concepts will be introduced in the following theoretical background chapters:

To start, chapter 2 will provide an overview on CSCL with a focus on co-present collaborative learning with touch interfaces. Next, chapter 3 will introduce theories of embodied cognition and how embodiment may be facilitated through the learning environment design. Chapter 4 will elaborate on the role of multifaceted processes during learning activities. Further, chapter 5 will provide more context regarding the learning domain proportional reasoning and the *Proportion* learning environment. Chapter 6 will discuss methodological considerations and techniques for researching children's co-present learning. Chapter 7 then will present the four studies done in the course of this thesis. To conclude, chapter 8 will provide an overall discussion of these four studies and suggest directions for future research.

2. CSCL – Computer-Supported Collaborative Learning

This first chapter will provide an overview on the overall context of the research done in this thesis: Peers learning together with technology. To this end, the concepts of collaborative learning with computers, and more specifically, co-present learning with one jointly used touch interface will be introduced.

Although learning may be conceptualized as an individual process of acquiring knowledge and conceptual understanding or skills, successful learning often happens in context, such as when peers come together to learn. This so-called *collaborative learning* allows peers to discuss or reflect together, so that the peers can learn from each other, be exposed do different perspectives, and potentially achieve deeper individual processing (Kirschner, Paas, & Kirschner, 2009; Stahl, 2005). If groups of learners manage to engage in meaningful interactions (see also chapter 4), collaborative learning can be more effective than individual learning (Dillenbourg, Järvelä, & Fischer, 2009). A common definition for collaborative learning is to characterize it as “interactions among peers” for the sake of learning (Dillenbourg et al., 2009, p. 3) or “two or more people working together toward a shared learning goal” (Jeong & Hmelo-Silver, 2016, p. 247). More specifically, the concept involves “a coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem” (Roschelle & Teasley, 1995, p. 70). Interactivity, i.e. learners influence each other’s reasoning (Dillenbourg, 1999), transactivity, i.e. learners’ contributions refer to each other and build up on each other (Weinberger, Hartmann, Schmitt, & Rummel, 2018), and co-construction of knowledge (Dillenbourg et al., 2009) are at the core of collaborative practices. Moreover, participants in collaborative learning groups are usually fairly symmetrical with respect to status and knowledge, as major differences in these factors would favor rather a tutoring (teacher – student) situation than a collaborative one (Dillenbourg, 1999; Mercer, 1996).

The learning processes in groups have been referred to as *group cognition* (Stahl, 2006). This group cognition encompasses cognitions on the group level – beyond individual cognitions – towards building knowledge together (Stahl, 2006, 2016). Argumentation – building, evaluating, and refining arguments and counterarguments – is a way of reasoning which unfolds its best potential in a discussion with a (learning) partner (Mercier, 2016). Consequently, this close interdependence between group members also entails that the learning outcomes (e.g., a product) are not necessarily ascribable to individual group members (Weinberger et al., 2018), as they are the result of a shared effort, and would need to be understood and evaluated on a group level (Stahl, 2016).

The concepts of *collaboration* and *cooperation* are partially used synonymously, but there are also attempts to untangle and define them (Dillenbourg, 1999). One prominent way of distinguishing them refers to how the learners divide the joint tasks: While collaborative learning implies that learners work closely on the same task, cooperative learning implies that learners work on sub-tasks individually and then combine the results as a final step (*horizontal* vs. *vertical division* of work; Dillenbourg, 1999). However, the line between cooperative and collaborative learning is not clear-cut (e.g., Schoor, Narciss, & Körndle, 2015). Learner prerequisites may change the interactions (Weinberger et al., 2018), and learners may also dynamically switch between collaborative and cooperative interactions (Jeong & Hmelo-Silver, 2016; Weinberger et al., 2018). A common core, however, is that collaborative and / or cooperative activities are characterized by a shared group goal that the group tries to achieve as a team (Jeong & Hmelo-Silver, 2016).

Collaborative learning with the help of technologies is commonly referred to as CSCL (computer-supported collaborative learning) (see also, Schmitt & Weinberger, 2018 / study I, section 7.1). CSCL scenarios can be classified as co-present or computer-mediated (Dillenbourg et al., 2009; Schmitt & Weinberger, 2018 / study I, section 7.1). In computer-mediated scenarios, learners typically interact via the internet (for example, synchronously in a chat, or asynchronously via forums) (Schmitt & Weinberger, 2018 / study I, section 7.1). In co-present scenarios, learners are physically and temporally together to learn collaboratively, either sharing one device or equipped with several individual devices (Schmitt & Weinberger, 2018 / study I, section 7.1). Co-present CSCL is the research focus in the present thesis.

Computational tools can present abstract domains in concrete ways, which may be conducive to mathematics learning (ter Vrugte et al., 2015; see also, Schmitt & Weinberger, 2019 / study III, section 7.3). By dynamically interacting with representations through these tools, learners are exposed to a wider range of learning opportunities (Ottmar & Landy, 2017). For example, in mathematics learning, a learner can actively engage with a dynamic software which provides feedback and discover the underlying mathematical rules and principles (Ottmar & Landy, 2017; Rick, 2012a; Schmitt & Weinberger, 2019 / study III, section 7.3). Displaying true-to-scale astronomic scales (Schneps et al., 2014) is another example of how technology can represent information in a way that would not be possible with traditional instructional means. In collaborative scenarios, a joint representation serves as an object that the learning partners can refer to and discuss about (Stahl, 2006).

Multi-touch devices, for instance tabletops (large interactive surfaces), allow several co-present learners to interact with a shared tool simultaneously, which is in contrast to early co-present scenarios where learners had to negotiate access to the input device (i.e., keyboard / mouse) (Schmitt & Weinberger, 2018 / study I, section 7.1). Touch interfaces are also named *natural user interfaces* (e.g., Black, Segal, Vitale, & Fadjo, 2012), and accordingly, as a consequence of the direct bodily interaction with the touch interface, the interaction is considered to be *natural* (Jamil et al., 2017).

Typically, co-present collaborative learning with multi-touch devices focused on shared tabletops. One example for collaborative learning on a shared tabletop is the *NumberNet* application, where children (10 years) collaboratively create diverse mathematical expressions for a specific target number (Mercier & Higgins, 2013). In other projects, young learners (11-15 years old) collaboratively create spider diagrams on a real-world problem such as sustainable energy (Jamil et al., 2017), or the learners (10-14 years old) collaboratively create a poster about their museum visit (Clayphan, Collins, Kay, Slawitschka, & Horder, 2018).

A small-in-size instance of interactive surface computers are tablets. Tablets, however, are often used as individual devices for individual learning, for example, pupils (10-11 years old) individually using iPads for mathematics learning (Carr, 2012). If tablets are used in collaborative scenarios, learners are often equipped with one tablet each rather than sharing one within a group. For instance, with the goal of representing a scientific phenomenon, children (10-11 years old) first create individual drawings on tablets, and then develop a joint drawing with a learning partner through a shared workspace, but still on separate tablets (Gijlers, Weinberger, van Dijk, Bollen, & van Joolingen, 2013). (It should be noted that the input in the study by Gijlers et al. (2013) was via an interactive pen, and not direct touch input). Early research on handheld devices articulated the danger that individually used handhelds, when every learner is equipped with their own device, impede interaction and collaboration (Liu, Chung, Chen, & Liu, 2009). However, an additional shared display promotes the occurrence of pointing gestures, productive discussions, shared visual focus, and collaborative interactions (Liu et al., 2009). Could the tablet (or handheld) device itself serve as this shared representation?

Rick (2012b) proposed the idea that one tablet could be used by multiple users simultaneously, i.e., tablets as "tiny tabletops" (para. 1). In such a scenario, the tablet could serve as the one joint representation for several co-present learners, just like in the case of tabletops. The *Proportion* app (e.g., Rick, 2012a), an app for fostering children's proportional reasoning, is an example for co-present CSCL featuring multi-touch interfaces. With such a jointly used multi-touch device, learners can possibly engage in rich interactions with a learning partner as well as the shared representation. In the present thesis, the *Proportion* app has been utilized in the studies II, III, and IV to investigate questions of co-present collaborative learning with a touch device. Specifically, how can researchers better survey children's experiences with a learning technology (Schmitt et al., 2019 / study II, section 7.2; see also section 6.1), how can the learners be supported to engage in meaningful learning processes (Schmitt & Weinberger, 2019 / study III, section 7.3; see also, section 4.3), and what is the role of learning process heterogeneity in such a scenario (Schmitt et al., in preparation / study IV, section 7.4; see also, section 4.4)?

The following chapter 3 will continue delineating the theoretical background of this thesis by providing an overview on theories of embodied cognition, and how these theories have been implemented in learning environment design.

3. Embodied Cognition and Learning Environment Design for Bodily Activities

Embodied cognition and the related terms *grounded cognition* and *embodied learning* refer to a perspective on human cognitive processing and learning that is attracting more and more attention. In a nutshell, this perspective points at the increasing evidence that the body, and bodily states and activities, play an important role for human cognition, and therefore also for learning.

The embodied perspective has been described and researched already for decades; as an example, Lakoff wrote in the 1980s: “Thought is embodied, that is, the structures used to put together our conceptual system grow out of bodily experiences and make sense in terms of it; moreover, the core of our conceptual system is directly grounded in perception, body movement, and experience of a physical and social character” (Lakoff, 1987, p. xiv). Theoretical accounts as well as empirical data supporting this perspective have been accumulated (and still are). In addition, technological advances make it possible to interact more richly with digital devices, beyond using keyboard and mouse. As a result, aspects of embodied learning, i.e., including the body more actively in learning scenarios, are becoming more popular in learning environment designs, with the goal of further developing the potential of (computer-supported) learning environments. Figure 1 demonstrates how *embodied cognition* and the related terms *grounded cognition* and *embodied learning* increasingly appear in the literature with an onset in the mid-90s (Google’s Ngram Viewer, Michel et al., 2011).

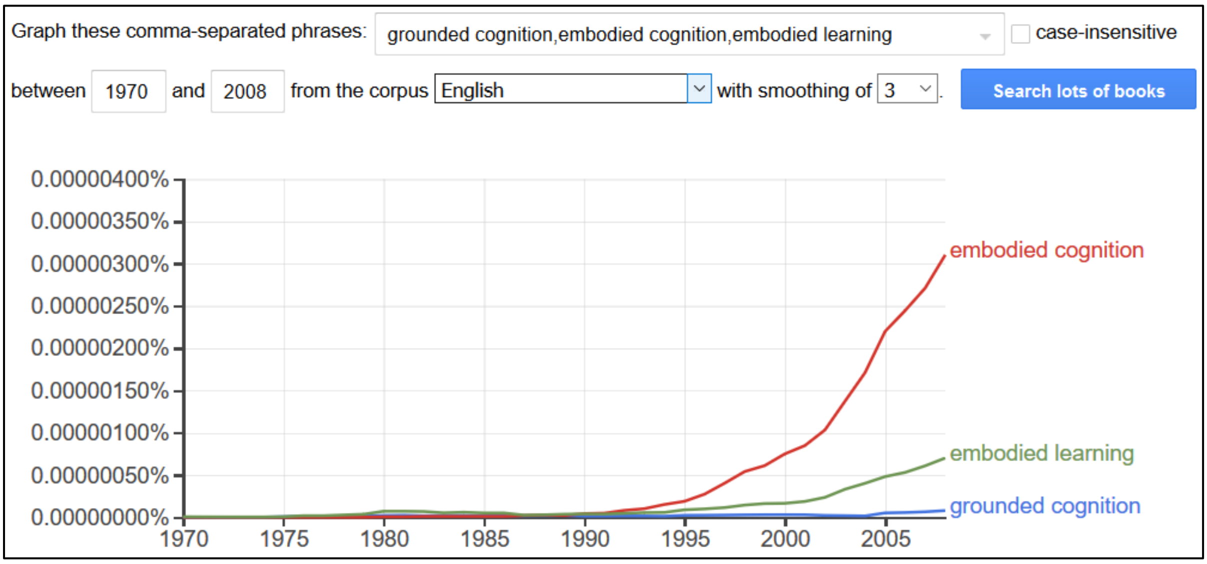


Figure 1. Embodied cognition and learning is gaining attraction in the literature. Screenshot from Google’s Ngram Viewer tool (https://books.google.com/ngrams), which provides data until the year 2008, Michel et al., 2011.

The *Proportion* learning environment, which is central to the present thesis' research, draws on embodied cognition and learning and engages the learner in bodily activities. Studies III and IV of this thesis will address open questions in the context of learning and learner interactions in embodied learning environments: Study III (Schmitt & Weinberger, 2019; section 7.3) investigates the question if bodily activities may need to be guided, or complemented with activities towards abstraction, for example, through verbal reflection. Study IV (Schmitt et al., in preparation; section 7.4) aims at untangling the role of bodily and cognitive processes, and their heterogeneity, in such an environment.

To theoretically ground the experimental research of studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4), this present chapter will give an overview on the concept of embodied cognition and its implications for learning. The chapter starts with defining the concept of embodied cognition and presenting the related main assumptions (section 3.1). In section 3.2, it follows an overview on the *Perceptual Symbol Systems (PSS)* theory, which aims to overcome accounts of amodal symbolic processing with a framework that deeply intertwines perception and cognition by proposing multimodal *perceptual symbols* (Barsalou, 1999). Next, in section 3.3, it will be elaborated on how the implications of theories of embodied cognition can be utilized to enhance learning processes. Finally, it will be illustrated how previous research has implemented ideas of embodied cognition in learning environment design (section 3.4), and specifically also in the case of learning with tablets (section 3.5).

3.1. Embodied Cognition: Definition and Main Assumptions

Are our thoughts, emotions, memories, or any other inner states and processes the result of a highly powerful computer (the brain) that receives input (e.g., a sensation on the skin), transforms it into amodal symbols (0 and 1 in the language of computers) and generates output as a response (e.g., the arm lifts up)? Why do we gesture when we speak to others or to ourselves? Is it true that standing upright with an open body posture makes us feel more self-confident? Is our conceptualization of ourselves and the world only a matter of input-output processing with the brain as the main (or sole) actor? Does it make sense to take children to the forest to see, smell, and touch the trees, or would showing them a picture have the same educational value? These and similar questions open up when we think about human information encoding, processing, storing and retrieval, and essentially also learning. According to the perspective of embodied cognition, “the mind must be understood in the context of its relationship to a physical body that interacts with the world" (Wilson, 2002, p. 625).

Both the terms *embodied cognition* and *grounded cognition* are used in prior research in similar ways (Barsalou, 2008) to refer to highly overlapping assumptions about cognition and the role of the body for human cognition. One distinction between these terms may be that *embodied* cognition lays the focus mostly on the relation of bodily states and bodily movements with cognition, whereas *grounded* cognition also includes cognition which is not directly coupled to momentary bodily states and movements, for example, mental simulations (Barsalou, 2008). However, also in literature that explicitly uses *embodied* cognition as the underlying theory, authors refer to simulations or imagined embodiment in absence of a stimulus or bodily activity (e.g., Black et al., 2012; Wilson, 2002).

Within these emerging frameworks of embodied human cognition, it is assumed that perception and action are closely coupled with cognition (Barsalou, 2008; Spackman & Yanchar, 2014), rejecting the computer metaphor of how the mind works (Niedenthal, Barsalou, Winkielman, Krauth-Gruber, & Ric, 2005). A computer senses input through an input device, and then translates this input into amodal symbols, and similar explanations had been stated for human processing of information as well (Barsalou, 1999), with the brain essentially doing computation like a computing device does (Curzon & McOwan, 2018; Edelmann, 2008). In this point of view, the brain and body constitute a hardware and the mind constitutes a software running on that hardware (Spackman & Yanchar, 2014), but the underlying hardware is somewhat arbitrary and could be easily replaced (Curzon & McOwan, 2018; Niedenthal et al., 2005), making cognitions disembodied. This perspective of the mind as an information processing entity has been referred to as *standard* or *traditional cognitive science* (Spackman & Yanchar, 2014).

From the embodied cognition perspective, there are two wide-spread assumptions about human cognition that are no longer supported (e.g., Barsalou, 1999):

(a) Embodied cognition theories reject the idea of a clear separation between the brain as the central control unit and the body as the entity to solely sense the environment and execute the orders of the brain (Barsalou, 1999; Niedenthal et al., 2005).
(b) Embodied cognition theories reject the idea of a translation from modal representations of perception, action, and introspection into an amodal symbolic representation (Barsalou, 1999, 2008).

Regarding the first assumption (a), non-embodied theories had assumed that there is a *recording system* and a separate *conceptual system* (Barsalou, 1999), with different tasks each: The recording system was assumed to gather holistic information in the environment without interpreting any of it, the conceptual system then was assumed to make sense of the recording, to draw inferences, to construct concepts, and to formulate propositions based on the information gathered with the recording system (Barsalou, 1999). From the embodied cognition perspective, however, there is no such thing as a private mind that can be separated from the body and our surroundings (Spackman & Yanchar, 2014).

Regarding the second assumption (b), both non-embodied and embodied accounts of cognition base themselves on the notion of (mental) internal representations (Spackman & Yanchar, 2014). Humans are capable of more than real-time bodily interactions with their immediate surroundings, they are also capable of abstraction and reasoning, which makes the assumption of some sort of mental representations necessary, that can be accessed and manipulated also in absence of a concrete stimulus (Spackman & Yanchar, 2014; Wilson, 2002). However, standard views assumed a re-description of concrete perceptual information into a symbolic amodal representation in a different neural system, which in turn serves as the basis for higher cognitive functions, like language and thinking processes (Barsalou, 1999). This point may be illustrated with an analogy: The word *chair* vs. the actual object of a chair (Barsalou, 1999). An amodal redescription of the perceptual representation of a bodily experienced chair is similar to the arbitrary mapping of the letters (symbols) *c-h-a-i-r* to their physical counterpart (Barsalou, 1999). While the notion of amodal symbolic representations used to be wide-spread and accepted, there is not much empirical evidence supporting it (Barsalou, 1999, 2008; Niedenthal et al., 2005). Glenberg (2010) argues that humans are "symbolic creatures" (p. 587) majorly dealing with symbolic representations in daily life (language, math), but that these symbols are grounded in our bodies through sensing / perceiving, emotionally experiencing, and interacting with our environment. Abstract symbols alone, without being associated with real-world counterparts, that we can engage with through our bodies, would be meaningless (Glenberg, 2010).

Theories of embodiment are diverse (Barsalou, 2008; Black et al., 2012; Glenberg, 2010; Spackman & Yanchar, 2014), but converge on the idea that "psychological processes are influenced by the body, including body morphology, sensory systems, and motor systems" (Glenberg, 2010, p. 586), while rejecting the notion of amodal symbols (Barsalou, 2008). Embodied cognition theories entail that cognitive processing cannot be reduced to an abstract and isolated brain activity where input from the body is received and output commands are issued as a response, but sensorimotor processing itself seems to be central for human cognition (Wilson, 2002). Bodily states, bodily activities, and mental simulations of these are deeply involved in human cognition (Barsalou, 2008), and cognitive functions, such as memory, language, or processing of abstract concepts are embodied (Spackman & Yanchar, 2014).

From a grounded cognition perspective (Barsalou, 2008), theories of embodied / grounded cognition draw on two basic assumptions about human cognition:

(a) The brain integrates perceptual, motor, and introspective states the body experiences during a given situation into a multimodal representation (Barsalou, 1999, 2008)
(b) These multimodal representations are accessed later-on to simulate that past situation when needed (e.g., remembering something) (Barsalou, 2008; Glenberg, 2010; Wilson, 2002)

Regarding the first assumption (a), there is evidence for multimodal representations of object properties in sensory and motor systems of the brain (Martin, 2007) with "considerable overlap in the neural circuitry supporting perceiving, acting on, and knowing about objects" (p. 27), which Martin (2007) sees as in support of notions of embodied cognition.

Regarding the second assumption (b), embodied processing is assumed to operate in two separable forms, with *online* (or *physical*) vs. *offline* (or *imagined*) embodiment. Online embodiment (Niedenthal, et al., 2005; Wilson, 2002), respectively physical embodiment (Black et al., 2012), refers to embodied processing during live interaction with the environment or during active manipulations with the body. But, in addition to the role of the body for processing during live activities, embodied processing also occurs in absence of a stimulus as a purely mental simulation of the experience, which has been termed offline embodiment (Niedenthal et al., 2005; Wilson, 2002), respectively imagined embodiment (Black et al., 2012). These sensorimotor simulations can internally recreate (aspects of) the previously experienced situation (Wilson, 2002), while not physically being in that situation (Barsalou, 1999).

Further approaches towards explaining the embodied character of cognition target the use of metaphors in human language, and the role of so-called mirror neurons. Regarding the first aspect, metaphors in language, Glenberg (2010) conceptualizes the use of metaphors as a way of representing abstract information from concrete bodily interactions with our environment, for example "(…) we talk about relationships as if they were journeys with beginnings, middles, ends, rocky parts, as well as smooth parts; and we talk about mood and emotion as having a spatial dimension as when a happy person is described as up or flying high and a sad person is down in the dumps" (Glenberg, 2010, p. 587), an idea put forward by George Lakoff

(Gallese & Lakoff, 2005; Lakoff, 1987; as cited in Glenberg, 2010). Mirror neurons are a second aspect which may serve as an explanation for embodied processing (Barsalou, 2008; Glenberg, 2010; Sapolsky, 2018). Mirror neurons refer to the observation that neuronal activity is very similar for performing a task oneself and monitoring someone else performing this task (Glenberg, 2010), respectively, mental states of others are simulated in our own brains to make sense of them, creating empathy and facilitating social coordination (Barsalou, 2008). However, the role of mirror neurons is somewhat controversial, and while they seem to be very relevant for observational learning processes, their role for empathy has been questioned (Sapolsky, 2018).

In sum, researchers in the context of embodied or grounded cognition emphasize the importance of (bodily) actions and the subsequent sensorimotor simulations of these actions (Barsalou, 1999, 2008; Glenberg, 2010; Wilson, 2002). However, embodied cognition theories do not only cover cases such as concrete physical manipulation of external objects. In addition, according to Niedenthal et al. (2005), also processing of social and emotional information is embodied. Affective and bodily states are tightly coupled and affect each other (Barsalou, 2008). Not only do emotions become evident through our bodies, for example through facial expressions (Ekman, 1992), but also, prior research showed that it may work the other way around as well: In a study by Strack, Martin, and Stepper (1988), making people hold a pen with their teeth, i.e. manipulating the facial muscles to form a smile, increased funniness ratings of a cartoon. A later study (Soussignan, 2002) found that this facial feedback effect specifically worked when the genuine Duchenne smile was induced. A recent meta-analysis on the facial feedback hypothesis based on 138 studies concluded that, while the effects are rather small and not always consistent, emotions can indeed be induced by manipulating facial expression (Coles, Larsen, & Lench, 2019). The role of emotional processes for learning is further elaborated in section 4.2.1.

A lot of the literature on embodied cognition remains fairly vague as to what exactly the underlying cognitive model looks like. Often, the claims do not go much beyond the idea that the body and its interactions with our environment are involved in cognition in some way. One notable exception is the PSS theory by Barsalou (1999). As Barsalou (1999) made a great effort of actually constituting a coherent model of human cognition with the PSS theory, an overview on this theory will be presented in the following section 3.2.

3.2. The Perceptual Symbol Systems (PSS) Theory

The main claim of the PSS (Perceptual Symbol Systems) theory could be condensed to: "(…) the dissociation between perception and cognition is artificial and invalid. Because perception and cognition share common neural systems, they function simultaneously in the same mechanisms and cannot be divorced" (Barsalou, 1999, p. 603). In the following paragraphs, I will give an overview on Barsalou's (1999) original work on the PSS theory:

The PSS theory assumes that a relevant subset of neural states during the perception of the environment and the own body is being represented as multimodal perceptual symbols (Barsalou, 1999). Depending on the modality of this perception (vision, hearing, haptic information, smell, taste, proprioception, or introspection), these differing types of information are represented in different associated brain areas (Barsalou, 1999). The perceptual symbols are mostly unconscious in nature, they do not require active imagery; however, selective attention towards the environment can play a role in making the experience more conscious in nature (Barsalou, 1999). Particularly for unconscious processing, the perceptual symbols can be componential, i.e., there is not necessarily a holistic representation of all aspects of a stimulus, but instead, there is very specific neuronal activation, for example, the shape of an object may be represented but its orientation may not be (Barsalou, 1999).

The PSS theory conceptualizes the perceptual symbols as dynamic, meaning that re-encountering similar environmental or bodily states may result in variations of the previous neural states. Re-activation of already existing perceptual symbols when re-encountering a stimulus may shape the perception and processing of that re-encountered stimulus (Barsalou, 1999). Furthermore, the perceptual symbols are not exclusively tied to specific individuals or instances of an object, but can also represent similar instances within the same broader category of information (Barsalou, 1999). The information stored in perceptual symbols is indeterminate and qualitative in nature, for instance, rather than representing an exact number of stripes, their length and thickness, the symbols can represent a stripe pattern (Barsalou, 1999).

Perceptual symbols with related information are being integrated across modalities and over time to form more complete representations of stimuli, a *frame* (Barsalou, 1999). At the heart of the PSS theory are the concepts of *simulators* and *simulations*. Simulation refers to processing in absence of a stimulus, and a simulator entails a frame plus its associated simulations (Barsalou, 1999). It is argued that "related [perceptual] symbols become organized into a simulator that allows the cognitive system to construct specific simulations of an entity or event in its absence" (Barsalou, 1999, p. 586). These simulations, however, are assumed to be most of the time incomplete and biased, due to mechanisms of selective attention, simplification, and extrapolation during encoding (Barsalou, 1999). Within the PSS framework, learning can be understood as the process of constructing simulators that in turn can produce appropriate simulations for the situations and objects humans have to deal with (Barsalou, 1999).

Pulling these ideas together, Barsalou (1999) argues against a modular view of strictly separated sensory-motor systems vs. amodal symbol systems in the brain, and claims that instead, perceptual symbols are directly present within the sensory-motor systems, making the notion of amodal symbol systems dispensable. The author argues that the perceptual symbol system is a "basic conceptual system, not a recording system" (Barsalou, 1999, p. 592). Thus, perception and cognition may operate within the same neural system, with perception being more than just an input generator for cognition based on amodal symbols, but a perceptual system may be representing knowledge and concepts itself (Barsalou, 1999). This perceptual symbol system enables the following points a – d, according to Barsalou (1999):

(a) *Productivity*: Going beyond the stored symbols and constructing complex representations also of situations or objects that are not being experienced, which allows imagination, as well as to share experiences between people through discussion
(b) *Propositions*: Adequately interpreting the situations and objects around us by mapping live perception with the concepts / simulators already acquired
(c) *Variable embodiment*: Being able to produce adapted perceptual symbols, depending on the individually encountered environments, situations, or objects
(d) *Abstract concepts*: Processes of retrieving a simulation relevant to the concept, including introspective states, in order to frame an abstract concept, and focusing on the most relevant part of that simulation are assumed to be central to represent an abstract concept

While testing or refining the individual aspects of the PSS theory is beyond the scope of this thesis, presenting one elaborated theory of how cognition may be embodied in more depth may help to better understand the context of research in embodied cognition or with embodied learning environments. Indeed, much of the prior research discussed in this thesis refers back to Barsalou's theory of perceptual symbol systems (Black et al., 2012; Cook et al., 2008; Glenberg, 2010; Niedenthal et al., 2005; Wilson, 2002).

Beyond the implications for the development of theories and models of the human mind and how we process, store, and retrieve information, including abstract concepts, embodied cognition theories are also relevant for understanding and supporting learning processes, and have therefore been influential for learning sciences research. The following sections 3.3, 3.4, and 3.5 will elaborate on how an embodied perspective has been applied to questions of learning and teaching.

3.3. Embodied Cognition for Learning

An embodied cognition perspective may contribute to advancing theoretical assumptions about how learners acquire knowledge and concepts; this necessitates the need to also regard and analyze non-verbal processes during learning activities (see also, sections 4.2 and 6.2). Further, the embodied cognition perspective brings up the question to what extent the design of instructions and learning environments would need to adopt an embodied perspective as well.

With an embodied cognition approach in mind, educational practices may change, and possibly improve, through the acknowledgment of the need to incorporate bodily experiences for enabling the grounding of abstract symbols, such as numbers (Glenberg, 2010). When people give explanations, they naturally use gestures that are related to their cognitions (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001). Bodily experiences may be facilitated through fostering bodily activities in a learning setting, and explicitly including body movements as part of the learning activities. This perspective of bodily learning is not necessarily new. For instance, already Vygotsky (1962) pointed out the role of direct and practical experiences for children's formation of concepts. Also the Montessori pedagogy, originally developed in the beginning of the 20th century (see Isaacs, 2018), includes active manipulation of objects involving multiple senses for fostering understanding of abstract concepts (Isaacs, 2018; Zuckerman, Arida, & Resnick, 2005).

The following paragraphs will provide examples of more recent research demonstrating how an embodied approach may support learning processes.

For example, embodiment (or lack of embodiment) may play a role for children's acquisition of reading skills (Glenberg, 2010). In reading, the letters are symbolic representations, and therefore, detached from the real-world actual counterparts, for instance, reading the letters of the written word *d-o-g* is not directly associated with the actual experience of encountering a real dog and seeing, hearing, touching, and interacting with it (Glenberg, 2010). This abstraction may make it difficult and demotivating for a child to learn reading, and children need support in establishing the connection (grounding) between the symbolic representation (the letters *d-o-g*) and the embodied / real-world representation (the actual dog) (Glenberg, 2010).

In mathematics learning, perceptual-motor activities may foster an understanding of mathematical symbols as actual physical objects that can be dynamically manipulated (Ottmar & Landy, 2017). This assumption is in line with the PSS theory (section 3.2), in the sense that formal symbol manipulation (e.g., in mathematics) can be facilitated by internally substituting the formal symbols with non-formal counterparts, e.g. real objects, to ease mental simulations (Barsalou, 1999). Black et al. (2012) describe this as a learning process in several steps: First, the learner needs to engage in an actual bodily experience, second, the learner needs to learn how to simulate this experience mentally without the live bodily aspect, in order to accomplish the third and last step, namely to reactivate this mental simulation when encountering formal symbols (Black et al., 2012). In the *Proportion* learning environment (e.g., Rick, 2012a) used in the present thesis, learners can actively manipulate a screen representation of a mathematical condition, for example, the proportional relation between two different integers.

Also the role of gestures in bodily interacting with a learning environment has been further untangled. Based on prior research indicating that gesturing and learning are related, Cook et al. (2008) were interested in determining a possible causal relationship of gesturing fostering learning (rather than gesturing only being an indicator for learning). In their study in the domain of mathematics, children (third and fourth grade) first observed an instructor who solved equations. The instructor used accompanying speech, or gestures, or both speech and gestures, while solving the equations. Next, the children repeated the instructor’s task, either in the speech condition (reproducing the instructor’s comments), in the gesture condition (reproducing a hand gesture the instructor was modelling), or the combined condition (reproducing speech and gestures), and solved further new tasks. While children of all three experimental conditions were equally successful in a post-test, only the children in the gesture conditions retained the knowledge four weeks later, indicating that involving the body in learning furthers lasting knowledge construction (Cook et al., 2008). However, not all gestures are suitable: For gestures to be beneficial for learning, it matters that they are congruent with the concepts to be learned (*gestural conceptual mapping*; Black et al., 2012). Learning environment design can facilitate the emergence of congruent gestures, for example, a tapping gesture would be congruent for the concept of discrete numbers, and a sliding gesture would be congruent for the concept of continuity (Black et al., 2012).

While it seems relatively clear that bodily activities are facilitative for learning processes, there are several different explanations of how and why this is the case. In the following paragraph, different perspectives on the role of bodily activity for leaning will be presented, primarily based on work done by Cook et al. (2008), Davidsen and Ryberg (2017), and Ottmar and Landy (2017):

(a) **Unburden working memory**: Using the body as another format of representing information may unburden working memory (Cook et al., 2008; Goldin-Meadow et al., 2001; Ottmar & Landy, 2017; Wilson, 2002). Davidsen and Ryberg (2017) speak of bodily resources as "cognitive auxiliary tools" (p. 82), which support knowledge building by enabling to store and represent information, for instance, with the fingers.

(b) **Facilitating information encoding**: The bodily activities themselves may serve processes of encoding information (Abrahamson, 2017; Cook et al., 2008; Niedenthal et al., 2005). Active manipulation or physical simulation, i.e., embodied experiences of a complex concept, may promote an intuitive understanding of this concept (Schneps et al., 2014), which may prepare the learner to better deal with the subsequently introduced symbolic formal representations related to the concept (Ottmar & Landy, 2017; Reinholz, Trninic, Howison, & Abrahamson, 2010).

(c) **The role of gestures**: Gesturing, for example, pointing, may serve the function of making learning-relevant features of the environment more salient, and to connect internal representations with the external world (Cook et al., 2008), or to facilitate information retrieval (Barsalou, 2008).

(d) **Uncovering strategies**: Bodily activities may lead to uncovering strategies which would not be evident in non-bodily activities, and these new strategies may be transferred to more traditional task formats and representations as well (Ottmar & Landy, 2017).

(e) **Triggering reflection**: The bodily experiences, which may be new and possibly illuminative for the learner, may result in reflective processes about these experiences (Ottmar & Landy, 2017).

(f) **Social perspective – Facilitating communication, illustration, and instruction**: Bodily movements and gestures help to better communicate and illustrate the conveyed information for the recipient (Barsalou, 2008; Davidsen & Ryberg, 2017), which is particularly important for social situations that require effective communication and collaboration. For instance, learners may perform a gesture to explain the concept of scale in a collaborative learning scenario (Davidsen & Ryberg, 2017). The bodily expressions also serve to coordinate interactions, e.g. pushing each other away, or granting space (Davidsen & Ryberg, 2017); behaviors which have been characterized as "dances between the participants" (Davidsen & Ryberg, 2017, p. 81). Also, bodily movements can have the function of *shepherding* in the sense that one peer instructs a learning partner of how to approach a task, for example, by moving the learning partner's hand (Davidsen & Ryberg, 2017). Finally, in a collaborative scenario, several learners together may also collectively embody concepts (Danish, Enyedy, Saleh, Lee, & Andrade, 2015; see also section 3.4).

Summarized, bodily activities seem to be a valuable source for facilitating learning. Specifically, in mathematics, connecting abstract symbolic representations to concrete hands-on manipulations may foster learners' understanding. The embodied approach to learning is relevant to the current CSCL research. Some researchers even proclaimed a "bodily turn in learning and within CSCL" (Davidsen & Ryberg, 2017, p. 67). Moreover, the *Computer Supported Collaborative Learning* conference in 2019 had *4E* learning and cognition as their main theme: "A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings" (Lund et al., 2019, p. ix). The following section (3.4) will give an overview on learning environments that explicitly integrate bodily activities as part of their design.

3.4. Embodied Learning Environments

Building on ideas of embodied cognition, researchers have started to design learning environments that explicitly integrate the body into the learning activities (Black et al., 2012). Particularly new technologies, for instance, tangible user interfaces, allow for uncovering scientific phenomena in a concrete *hands-on* way, making the underlying principles of these phenomena visible or graspable (Sakr, Jewitt, & Price, 2014). Therefore, new technologies may be suitable for furthering embodied cognition and learning, and have in fact been used to create learning environments with a focus on bodily involvement for carrying out the learning activities. One way of classifying embodied learning activities is through the *Instructional Embodiment Framework* by Black et al. (2012): Within this framework, on can distinguish between *physical* and *imagined* embodiment. This distinction is analogous to the one between on- and offline embodiment (Niedenthal et al., 2005; Wilson, 2002; see section 3.1). In the case of physical embodiment, the learner is actively moving their own body (*direct embodiment*), or is controlling the movements of an avatar (*surrogate* and *augmented embodiment*) (Black et al., 2012). In the case of imagined embodiment, the learner is mentally simulating the embodied activity, which may be *explicit* or *implicit* (Black et al., 2012). The following paragraphs will present two different embodied learning environments that focus on learning in the domains of natural sciences and mathematics.

The first example is a learning environment for mathematics learning, and specifically targets children's difficulties in proportional reasoning (see also section 5.1). Can proportional relations be experienced through the body? With the *Mathematical Image Trainer*, young learners (grades four to six) work on their proportional reasoning abilities through senso-motorical experiences (Reinholz et al., 2010). In this learning scenario, the learners control two virtual objects (displayed on a screen) with two handheld devices that they move in the air with the goal of representing proportional relations. The learners receive feedback, i.e., the screen becomes red, yellow, or green, depending on how close the learners are to the solution (the target proportion, for instance, 1 : 3). During the course of the learning activity, the screen display is being enriched, first with grids, and second with numerical labels, to better connect the bodily experience to the learners' arithmetic knowledge. However, the target proportion is not made explicit to the learners and they are only instructed to "make the screen green" (Reinholz et al., 2010, p. 1488). This approach is supposed to facilitate learners' construction of an embodied representation of how a given proportion (e.g., 1 : 3) can be realized, namely by an indefinite number of different hand configurations as long as their relation corresponds to the target proportion. Reinholz et al. (2010) report that the students benefitted from the embodied experience, as they were able to articulate observations relevant for domain understanding in interviews with the researchers. Reinholz et al. (2010) explain it like this (direct quote):

> Students' successful modeling of a problem situation even before engaging normative mathematical media, notation systems, and formats, suggests promise in instructional designs affording embodied, presymbolic quantitative reasoning. Such experiences enable students to struggle with qualitative aspects of a mathematical domain and discover its quantitative principles before they are burdened with the supplementary cognitive load of the disciplinarily requisite inscription and calculation procedures. Namely, as students are guided through a sequence of insights into the properties of a mathematical phenomenon, they can perform core conceptual work even prior to symbolic articulation. (pp. 1494-1495)

According to the Instructional Embodiment Framework (Black et al., 2012), the Mathematical Image Trainer environment may be classified as direct physical embodiment.

Not only can learners embody concepts individually, like in the example above, but there are also learning environments in which learners collectively embody the content to be acquired: In the *Science through Technology Enhanced Play (STEP)* project (Danish et al., 2015), children aged between seven and eight years can experience scientific phenomena with mixed-reality tools (combining real-world with virtual representations): For instance, the participants act out water particles in their different aggregate states. They do so by moving around the classroom as a group while their movement is being tracked and transformed into a computer simulation displayed on a large screen. Standing still and close to each other changes the aggregate state to ice, moving around very vividly changes the aggregate state to gas. Within the Instructional Embodiment Framework (Black et al., 2012), this STEP scenario may be classified as augmented direct embodiment. Danish et al. (2015) found their intervention to foster learning about this domain (measured through pre- / post-interviews), but also that the students would sometimes get lost in action and display overly playful behavior (e.g., playing tag). Furthermore, the learners benefitted from a teacher who prompted a reflective discussion about the activity, redirecting the focus from the purely physical aspects of moving around towards the underlying concepts (Danish et al., 2015).

These two examples show how learners can acquire concepts through bodily activities. However, not all embodied learning environments may require full bodily involvement with learners engaging in large movements. The bodily involvement may also come on smaller scale, for example with touch screens, which will be introduced in the following section.

3.5. Tablets as Embodied Learning Environments

Also interactive touch interfaces, for instance tablets, afford active involvement with the body through touch and bodily activities, therefore, may benefit embodied processing and learning (Black et al., 2012). The following paragraphs will exemplify this approach of embodied learning with tablets with two studies:

The first presented example follows the idea that active manipulation of screen representations via touch may be conducive to mathematics learning: Ottmar and Landy (2017) introduce the approach of *Pushing Symbols* to facilitate mathematics learning by involving so-called “perceptual-motor routines” (p. 53) on tablets. This approach implies that in addition to formal manipulation of formal mathematical notations, learning activities should include perceptual-motor activities, for example gestures (Ottmar & Landy, 2017). In their study, Ottmar and Landy (2017) compared two ways of combining formal with perceptual-motor learning activities, with the learning goal of being able to simplify mathematical expressions. They argue that, while traditionally, learners encounter the formal representations first and only afterwards get the chance to practice and experience them (Nathan, 2012; as cited in Ottmar & Landy, 2017), it may be conducive to learning to start with concrete perceptual-motor activities, and introduce the formalisms afterwards. The perceptual-motor activity (*dynamic lesson*) had learners actively manipulate mathematical objects via direct touch on tablets; the formalized activity (*static lesson*) had learners work with a static worked examples tablet program via a stylus input on tablets (Ottmar & Landy, 2017). As hypothesized, the learners (seventh graders) benefitted from initial perceptual-motor activities followed by a formal activity more than in the learning condition with the more traditional order (working with static formalisms first, followed by dynamic learning activities). Ottmar and Landy (2017) interpret this finding as follows: The concrete dynamic activity may offload the working memory as the information is represented through the body, enabling processing of new content (see also, point (a) in section 3.3). Also, the learners may encounter useful strategies that they are not able to extract from a static representation, but that help to engage with the static representations (see also, point (d) in section 3.3). Finally, “a perceptual routine itself may serve as the target of reasoning” (Ottmar & Landy, 2017, p. 71), in the sense that the active experiencing may trigger productive reflection processes (see also, point (e) in section 3.3).

Embodied learning with tablets may also be suitable to address misconceptions, such as learners’ nonscientific ideas about astronomic scales, by enabling a more true-to-reality sort of experience, which may be difficult to provide with traditional learning materials. Traditionally, representations of the solar system often displayed exaggerated versions of the planetary bodies (planets look much bigger and closer together as they actually are), due to limited space on paper or a screen (Schneps et al., 2014). With the *Solar Walk* tablet app, learners can experience true-to-scale sizes and distances by dynamically exploring the solar system and zooming in and out (Schneps et al., 2014). In their study, the participants (around 15 years old) explored the app either in a *true-to-scale* mode or in an *orrery* mode (traditional exaggerated representation). Schneps et al. (2014) found that the learners in the true-to-scale condition were more likely to overcome their misconceptions about astronomic scales than the learners in the orrery condition. The authors assume that this effect may be due to the greater affordance for gestures (and, therefore, more embodiment) in the true-to-scale mode, as in this mode, the learners had to zoom in and out quite a lot, in order to explore the solar system with (accurately represented) very small and wide-spread planetary bodies. While a better understanding of astronomic scales could be achieved through the embodied activity, there was no gain in other concepts such as the origin of seasons, which led Schneps et al. (2014) to conclude that additional scaffolding may be needed to explicitly address learners' misconceptions.

Both of these environments (Ottmar & Landy, 2017; Schneps et al., 2014) may be classified as targeting direct physical embodiment according to the Instructional Embodiment Framework (Black et al., 2012). Although tablets may not involve gross body movements as in the case of, for instance, the STEP activities (Danish et al., 2015), also tablets do enable a certain level of bodily involvement, which may grant learners better access to difficult topics in mathematics or natural sciences. Tablets as a wide-spread and relatively affordable technology may feasibly be integrated into formal and informal learning settings (see also, Rick, 2012a) and complement traditional learning and teaching approaches with the opportunity for embodied learning experiences.

3.6. Conclusion

Whether or not all cognitive processing is embodied, it seems to make sense to design embodied learning environments which draw on embodied cognition and afford bodily activities. Explicitly integrating bodily movements and facilitating hands-on experiences showed to be useful in advancing learners' understanding of a domain, respectively in addressing their misconceptions (e.g., Ottmar & Landy, 2017; Reinholz et al., 2010; Schneps et al., 2014).

While hands-on learning, i.e., concrete experiences with the environment through our bodies, maybe is the oldest form of any human (or non-human) learning, the combination with recent technological tools is an advancement. As Sakr et al. (2014) argue, the combination of bodily activity plus the digital representation is crucial. Technologies can represent information and properties that would be difficult to represent without technologies. To illustrate, in the previously presented project by Danish et al. (2015), the children may have also been able to act out the particles without any technological support. However, the live feedback of their activity, the matter of state, is only made possible through the tracking of their movement and the resulting computer simulation. Another advantage of technologies is that they allow for putting the physical activities on hold in order to engage in abstraction at any time, while the physically constructed representation remains visible (Sakr et al., 2014). Thus, in addition to the embodied experiences, a (digital) representation stores and displays information and can serve as a starting point for abstraction and reasoning processes.

In the present thesis, the embodied learning environment *Proportion* (e.g., Rick, 2012a; see chapter 5) running on multi-touch tablets is being utilized to help students acquiring a better understanding of proportional reasoning by bodily experiencing and manipulating proportional relations. Also in the *Proportion* learning environment, feedback about the proportional relation between two numbers can only be realized technologically, which makes the technology a crucial part of the hands-on activity. Further, the screen may serve as a shared representation of the task and enable joint reflection of the bodily experiences.

Indeed, in spite of the encouraging findings of learning in embodied environments presented in sections 3.4 and 3.5, it is also reported that bodily activity alone may not be sufficient for acquiring a good conceptual understanding, and that additional (verbal) reflection on these bodily learning activities may be a highly relevant factor for successful learning (Danish et al., 2015; Schneps et al., 2014). Such a reflective activity may help learners to draw connections between the bodily experience and a (seemingly) abstract underlying concept, typically represented through formal symbols, for example, numbers in mathematics. Also Duschl (2008) argues that education practices need to strive towards a productive combination of practical hands-on activities with explicit discussion of the underlying principles to develop conceptual knowledge. However, while these reflective activities may be supportive for learners' understanding, they may also have a disruptive effect in learning scenarios that are all about engaging with content through the body. The effects of supplementing an embodied learning environment with additional prompts for strategic behavior and reflective dialogue are reported in study III (Schmitt & Weinberger, 2019; section 7.3).

Further, it remains unclear how exactly participants' learning processes on different dimensions unfold while they immerse themselves with their body in an embodied learning environment. Specifically in the case of collaborative learning, learners with homogeneous or heterogeneous learning processes may be grouped together in a setting where they interact not only verbally, but also with their whole bodies. The effects of learning process heterogeneity on two dimensions (bodily and cognitive) are reported in study IV (Schmitt et al., in preparation; section 7.4).

The following chapter 4 will proceed with giving an overview on verbal and non-verbal learning processes which are relevant to learning and the work done in this thesis.

4. Multifaceted Learning Processes and their Significance for Learning

For a thorough understanding of the differential effects of different learning environment designs, it is crucial to investigate the learning processes in addition to the learning outcomes. An interest towards multiple qualities of learners' collaboration, for example, how they express emotions, or the nature of their talk, necessitates an analysis of learners' ongoing learning processes on multiple dimensions. Different to computer-mediated CSCL, embodied co-present learning with a rich interactive technology involves bodily activities of several learners (Liu et al., 2009; see also, Schmitt & Weinberger, 2018 / study I, section 7.1), producing interactions on multiple dimensions. For instance, in addition to verbal contributions, learners may also express different emotions through their body or display different degrees of task focused behavior. These learning-relevant aspects of co-present learning processes would get lost if one would only analyze learning outcomes, or survey these aspects in a retrospective way (Sakr, Jewitt, & Price, 2016; see also, Liu et al., 2009). In contrast, a multimodal process-oriented approach may be suitable for capturing these learning processes. A multimodal analysis can, for example, include attention towards speech, body movements, and interaction with the learning technology (Sakr et al., 2016), or talk, visual focus, and gestures (Liu et al., 2009). In the present thesis, specifically in studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4), task focus as a basic prerequisite for learning, as well as the bodily expression of emotions, and the quality of learners' verbal contributions are analyzed (for the analysis methods, see section 6.2).

The following sections will give an overview on research related to these process variables, and motivate their inclusion in the present thesis, i.e., discuss their significance for learning and collaborative processes. Section 4.1 will elaborate on the importance of learners' verbal contributions, and introduce two relevant dimensions of their talk: Transactivity (section 4.1.1), and epistemic quality (section 4.1.2). Section 4.2 will shift towards non-verbal facets of learners' processes, and specifically include the aspects of emotions (section 4.2.1), and task focus (section 4.2.2).

4.1. Verbal Learning Processes: Transactivity and Epistemic Quality

Dialogue between humans can be understood as a "social mode of thinking" (Mercer, 2004, p. 141), with high quality dialogue being central to human learning (Mercer, 2004). Learners' discourse is not only a means to the end of a learning product, but it is learning and knowledge construction in itself (Pontecorvo & Girardet, 1993). Although communication and interaction are not exclusively dependent on language in co-present collaborative learning scenarios, learners' dialogue is still a central feature of their interactive processes and often a focus of research, when one is interested in the quality of the collaborative learning processes (e.g., Gijlers et al., 2013; Roschelle & Teasley, 1995). Learners' talk enables the sharing and constructing of knowledge (Liu et al., 2009; Mercer, 1996; Stahl, 2006), and makes learners' cognitive processes and joint reasoning visible (Mercer, 1996; Stahl, 2006). Moreover, the talk between peers can facilitate generalization of understanding, enabling learners to make a successful transfer from one concrete problem towards the underlying principles of this problem, and to solve related problems (Mercer, 1996). Already children are capable of engaging in complex reasoning together and constructing knowledge and theories through discourse (Duschl, 2008; Pontecorvo & Girardet, 1993). In studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4) of this thesis, talk – in addition to the bodily activities – is regarded as an important factor for learners' knowledge construction. In general, a distinction can be made between the *content space* and the *relational space* in learners' discussion (Janssen, Erkens, & Kirschner, 2011). While the first sort of talk refers to moves towards solving, coordinating and regulating the learning tasks, the latter sort of talk deals with the social aspects of the collaborative learning situation, such as coordination of responsibilities, or meta-talk about the success of the group project (Janssen et al., 2011).

But how does talk enable learning, and what are constituting features of high quality talk? In collaborative learning research, it is assumed that talk, and the quality and frequency of certain types of talk, relate to and possibly also promote learning outcomes (Baker & Lund, 1997; Weinberger & Fischer, 2006). For example, it is crucial that learners are able to express themselves clearly (Mercer, 1996), respectively, to verbalize their own understanding of a concept (van Boxtel, van der Linden, & Kanselaar, 2000), and that learners ask questions and respond to their peers' questions (van Boxtel et al., 2000). Based on a body of empirical research, Webb (1989) presents a model of learners' talk types and how these talk types lead to other sorts of talk types and finally different qualities of learning outcomes. For example, according to the model presented in Webb (1989), a student posing a high-level question and then receiving a high-level response is likely to succeed in the problem solving activities, and to elaborate on a high level as well, resulting in a high understanding. In contrast, a student posing a low-level question and subsequently receiving none or only poorly elaborated feedback will likely struggle in the problem solving activities, with very limited learning achievements as a result (see the model presented in Webb, 1989).

So, the extent to which learners manage to formulate elaborated questions and explanations seems to matter for learning. According to Baker and Lund (1997), "reflective interactions (…) involve explanation, justification and evaluation" (p. 176). Giving explanations is positively correlated with learning outcomes for the explaining learner, possibly due to processes of cognitive restructuring in order to be able to give a good explanation, which may further the own understanding, respectively make a potential lack of understanding salient (Webb, 1991). While it may seem plausible that receiving explanations by a learning partner is a useful source of information, empirical findings do not generally support this assumption: Webb (1991) reasons that the explanations may often not be targeting the specific problems, or may just not be understood by the recipient, or may even take away a sense of responsibility to try to understand the problem oneself. In conclusion, giving explanations is beneficial in collaborative learning, and should therefore be facilitated and prompted in a learning environment, but attention needs to be paid to the recipient's current needs (Webb, 1991).

Beyond giving explanations, the benefits of talk between peers are also due to the emergence of *socio-cognitive conflicts* – a learning opportunity in which learners manage to resolve a conflict in their shared understanding by restructuring their individual understanding (Mercer, 1996; King, 1990; Roschelle & Teasley, 1995; van Boxtel et al., 2000; Weinberger, Ertl, Fischer, & Mandl, 2005). In essence, productive talk involves joint reasoning, i.e., processes of analyzing, explaining, and drawing conclusions together (Mercer, 1996). From a grounded cognition perspective, being able to communicate through language enables humans to share their own mental simulations or influence their peers' simulations, as basic premise for coordination and collaboration (Barsalou, 1999).

To assess when learners' talk is more or less productive, researchers have developed different ways to classify the talk. For instance, Mercer (1996), who observed 9-10 year old children solving problems together at a computer, distinguishes between *disputational*, *cumulative* and *explorative* talk: Disputational talk is characterized by learners exchanging individual short statements that may be in disagreement, but without attempts to integrate the conflicting information, the nature of the talk is rather competitive (Mercer, 1996). Cumulative talk is characterized by learners accumulating individual pieces of knowledge by repeating and adding them together, but again without a critical stance of challenging each other, avoiding conflict (Mercer, 1996). Explorative talk is characterized by a critical stance towards the peer's contributions, by asking and responding to questions, by attempts to productively challenge the other's contributions and to engage in actual reasoning processes (Mercer, 1996). While disputational talk results in individual conclusions, cumulative and explorative talk aim for consensus, and particularly in explorative talk, the joint conclusion reflects the shared understanding that has been collaboratively built through the discussion (Mercer, 1996).

Beyond the talk types presented in Mercer (1996), there are many other ways of disentangling different facts of learners' talk. For example, one can identify several dimensions of a collaborative discussion ("specific discourse activities", Weinberger & Fischer, 2006, p. 72) and relate them to learning outcomes (Weinberger & Fischer, 2006). For instance, on the social side of learners' talk, learners may more or less build up on each other's contributions, and their epistemic discourse activities – referring to the content of the talk – can demonstrate the degree of learners' knowledge acquisition and application (Weinberger & Fischer, 2006). This question of disentangling facets of learners' talk is further elaborated on in section 6.2.3, with a detailed overview on ways to analyze learners' talk, along with the coding schemes that were developed and utilized in the course of the present thesis, specifically in studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4).

The present thesis regards the quality of learners' talk on two main dimensions: *How* and *what* (Weinberger et al., 2005). The *how* reflects the social level of learners' discussion (also called *transactivity*), and the *what* reflects the contents learners discuss (also called *epistemic activities / quality*) (Weinberger et al., 2005). The following sections 4.1.1 and 4.1.2 will elaborate on these two aspects of learners' talk, which showed to be particularly relevant for learning, and can indicate the quality of the learning processes.

4.1.1. The relevance of transactivity in dialogue

"They took alternate turns: but then so do opponents in tennis" (Mercer, 1996, p. 366). Learners may – on the surface – engage in dialogue, indicated by frequently taking turns with both learners being involved in verbalizations. However, this does not necessarily imply that the learners actually talk *with* each other. Turn-taking in dialogue needs to be cooperative in nature and enable the sharing of individual problem representations (Roschelle & Teasley, 1995). A measure of the amount that learners talk *with* each other is called *transactivity*. Two aspects are central to the concept of transactivity: First, learners make their knowledge explicit, for instance, by verbalizing their reasoning, and second, learners are willing to draw connections between their individual reasoning (Gweon, Jain, McDonough, Raj, & Rosé, 2013). In other words, transactivity refers to the amount that learners build on their learning partner's contributions, refer to these contributions, acknowledge, paraphrase, extend, complement, integrate, or challenge them (Gweon et al., 2013; Noroozi, Teasley, Biemans, Weinberger, Mulder, 2013; Roschelle & Teasley, 1995; Stahl, 2013; Teasley, 1997; Vogel et al., 2013;

Weinberger & Fischer, 2006). In contrast, low-transactive talk may simply consist of uncritical and quick consensus building rather than truly and critically engaging with the learning partner's contributions (Weinberger et al., 2005). Moreover, it has been found that learners influence each other with respect to the degree of transactivity (Vogel et al., 2013).

Transactive utterances can be on one hand rather conflict-oriented, as in productive cognitive conflict, or on the other hand rather agreement-oriented, as in completing and integrating the learning partner's contributions (Teasley, 1997). According to Berkowitz and Gibbs (1983), in a transactive dialogue, learners confront each other with their individual reasoning, and not just accumulate their viewpoints. They distinguish between transactive moves that are intended to "elicit or *re*-present another's reasoning" (p. 403), and transactive moves that actually aim at "operating on or transforming" (p. 403) the peer's reasoning (Berkowitz & Gibbs, 1983). The authors consider operating on and transforming the peer's verbalized knowledge as highly transactive, even the "truest forms of transacts" (Berkowitz & Gibbs, 1983, p. 403).

Learners challenging their learning partners by asking critical questions may foster elaboration of the learning material, and therefore, facilitate acquisition of knowledge (Weinberger et al., 2005). Thus, the degree of transactivity is a relevant aspect for the quality of collaborative talk and learning outcomes (Noroozi et al., 2013; Teasley, 1997). The following paragraphs will illustrate this with three studies in the context of collaborative learning in mathematics and natural sciences:

In the first study, Vogel et al. (2013) investigated the effects of scripts and heuristic worked examples on learning and transactive dialogue. They found that the combination of both means of instructional support was most successful for fostering transactivity, and that transactivity was a mediator for individual learning outcomes (Vogel et al., 2013). However, this important role of transactivity was only present in the case of self-generated transactive moves, whereas the transactive moves of the partner were not relevant for individual knowledge acquisition, underlining the importance that the learners themselves have to deeply engage with the partners' contributions for enhanced learning (Vogel et al., 2013).

In the second study, Blanton and Stylianou (2014) examined the relationships between transactivity in classroom discussions and skill development regarding mathematical proofs, using a qualitative sociocultural approach. They found that, over the time of several months, the students adopted a more transactive mode of discussing, and that this increased transactivity level helped to develop mathematical proof abilities (Blanton & Stylianou, 2014).

Finally, in a study with children drawing scientific phenomena, Gijlers et al. (2013) conclude, based on a regression analysis, that the degree of transactivity in dialogue is relevant for knowledge outcomes, specifically regarding learners' ability to explain the scientific phenomena, which underlines the importance of transactivity for the construction of complex knowledge structures (Gijlers et al., 2013).

To sum up, the degree of transactivity regards *how* learners talk to each other, to what extent they refer to each other (e.g., Berkowitz & Gibbs, 1983), and possibly engage in productive socio-cognitive conflicts (e.g., Noroozi et al., 2013), with higher degrees of transactivity being helpful for refining and building their understanding (e.g., Teasley, 1997). As such, it was considered relevant to include this important quality of learners' verbal interactions in studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4). The way how transactivity was exactly conceptualized and represented in a coding scheme in this thesis can be found in section 6.2.3.

4.1.2. The relevance of epistemic quality in dialogue

In addition to the extent that learners' refer to each other's contributions and productively challenge them, also the *content* of their contributions (the *what*; Weinberger et al., 2005) needs to be regarded, which is referred to as the *epistemic* dimension of talk (Weinberger & Fischer, 2006), and often centers around the nature of learners' explanations (Pontecorvo & Girardet, 1993). On an epistemic level, it first of all matters that learners engage in on-task talk (Weinberger & Fischer, 2006). Second, it is crucial that the talk content facilitates activities towards solving the learning tasks (Roschelle & Teasley, 1995; Tsovaltzi, Judele, Puhl, Weinberger, 2017), and towards elaborating on domain-relevant concepts and drawing connections between these concepts (Tsovaltzi et al., 2017). One can distinguish different epistemic activities, like discussing the task at hand, vs. discussing related theories more abstractly, vs. relating the task at hand with prior knowledge or newly acquired theoretical assumptions (Weinberger & Fischer, 2006). The different epistemic activities are relevant for knowledge acquisition. For example, while discussing the task at hand is important for solving the current task, it might not suffice to make progress in learning, as no new knowledge is being generated (Weinberger & Fischer, 2006). Particularly relevant for knowledge acquisition and application is to draw connections between the concrete problem and the related theories (Weinberger & Fischer, 2006). In sum, what matters is that learners engage in productive epistemic activities which facilitate knowledge construction (Weinberger et al., 2005).

However, beyond the plausible theoretical assumption that high epistemic quality in discussions is conducive to learning, there are not many studies empirically testing the relationship between epistemic quality and learning outcomes (e.g., measured with a formal test). One example is a study by Tsovaltzi et al. (2017), which found that, when learners are supported with an argumentation script, the level of elaboration of arguments (epistemic dimension) is a mediator for individual knowledge outcomes as well as knowledge convergence between learning partners. Beyond this, the lack of studies may be explained by the notion that epistemic quality in learners' discussion is a learning outcome in itself already, serving as an indicator for the level of collaborative knowledge co-construction (see also, Pontecorvo & Girardet, 1993).

In terms of the epistemic aspect of learning processes, two concepts closely related to epistemic quality – namely *epistemic curiosity* and *epistemic beliefs* – have been researched more extensively and will be shortly introduced: For instance, it was found that epistemic curiosity mediates the relationship between personality traits, specifically conscientiousness and openness to experience, and learning outcomes (Hassan, Bashir, & Mussel, 2015); a finding which makes researchers argue for trainings to enhance degrees of epistemic curiosity to facilitate learning (Hassan et al., 2015). As for epistemic beliefs, there is evidence that the nature of learners' epistemic beliefs predicts learners' self-reported emotions (Trevors, Muis, Pekrun, Sinatra, & Muijselaar, 2017). Specifically, a good fit between the epistemic framing of the task, for example complexity and certainty of knowledge, with learners' own epistemic beliefs predicted positive emotions (Trevors et al., 2017), and learners' emotions predicted learning outcomes, and mediated the relationship between epistemic beliefs and learning (Trevors et al., 2017). For example, presenting a text with conflicting information to a learner with matching epistemic beliefs (knowledge is complex and sometimes contradictory), fostered the emotions of surprise (neutral) and curiosity (positive), and these emotions in turn helped to memorize the contents (Trevors et al., 2017). Especially regarding epistemic beliefs, there is a wide body of research available (e.g., Franco et al., 2012; Mason, Ariasi, Boldrin, 2011; Tsai, Ning, Ho, Liang, & Lin, 2011), but as this concept is not central for this thesis, it will not be further elaborated on.

To sum up, the dimension of epistemic quality in learners' talk regards the contents of the talk, how learners include concepts and theories and draw connections between them (e.g., Weinberger & Fischer, 2006). Epistemic quality is often seen as central for fostering learning outcomes, even though this relationship is mostly argued for based on theoretical assumptions. Beyond a means to an end for learning outcomes, epistemic quality can also be thought of as a learning goal in itself (see also, Pontecorvo & Girardet, 1993), as an indicator for the quality of collaborative processes. Next to transactivity, also epistemic quality has been included as a central verbal process variable in studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4). The way how epistemic quality was exactly conceptualized and represented in a coding scheme in this thesis is described in section 6.2.3.

4.2. Non-Verbal Learning Processes: Emotions and Task Focus

“Language, however, does not occur in a vacuum” (Roschelle & Teasley, 1995, p. 95). Talk is not the only means in collaborative learning for constructing (shared) knowledge and understanding (Roschelle & Teasley, 1995), and “one major role of the computer in supporting collaborative learning is providing a context for the production of action and gesture” (Roschelle & Teasley, 1995, p. 78). Not everything learners know can be expressed with words alone, sometimes learners may utilize gestures, objects, or a shared representation, for instance a computer screen, to explain their understanding to their learning partner (Roschelle & Teasley, 1995). In conclusion, attention needs to be paid to learners’ non-verbal or bodily activities during their interactions with each other, for instance, their visual focus, gestures or bodily expression of emotion, in order to better understand their communicative, collaborative, and emotional processes, and ultimately learning (Chung, Lee, & Liu, 2013; Davidsen & Ryberg, 2017; Sakr et al., 2016).

The following sections 4.2.1 and 4.2.2 will introduce two aspects of learners’ non-verbal processes, which are salient in co-present CSCL scenarios and, at the same time, relevant for learning: Learners’ emotions and task focus.

4.2.1. The relevance of emotional processes

What are emotions, and specifically, what are emotions in a learning context? Generally, emotions can be defined as the responses that develop in relation to events that someone may interpret as supporting or hindering the achievement of individually relevant goals (Frijda, 1988). Emotions in the context of a learning situation are conceptualized in diverse ways in the literature, for example, as *academic emotions* (Pekrun, Goetz, Titz, & Perry, 2002), *achievement emotions*, or *epistemic emotions* (Muis et al., 2015). While achievement emotions relate to emotions – for instance, frustration – in response to an achievement or perceived lack of achievement, epistemic emotions may occur in response to an epistemic or cognitive conflict, for example, dealing with multiple solutions or conflicting information (Muis et al., 2015). Other than that, research in the context of learners’ emotions often draws on Ekman’s

(1992) basic emotions (e.g., in Lehman, Matthews, D'Mello, & Person, 2008), or distinguishes two main branches of emotions: On one hand *positive emotions* like excitement, happiness, enjoyment, hope, pride, relief, or curiosity, and on the other hand, *negative emotions* like frustration, irritation, anxiety, confusion, hopelessness, disappointment, sadness, anger, or boredom (Deater-Deckard, Chang, & Evans, 2013; Knörzer, Brünken, & Park, 2016; Muis et al., 2015; Pekrun et al., 2002; Trevors et al., 2017). Moreover, in collaborative settings, emotions may not only be related to the task or the self, but may also originate from the interactions with the group members (*social emotions*) and the general learning environment (Järvenoja & Järvelä, 2005).

Emotions can be linked to learning outcomes through the assumption that they impact learners' cognitive processes (Pekrun et al., 2002). It has been found that positive emotions correlate with a broad variety of beneficial variables, such as interest, motivation, effort, metacognitive strategies, or critical thinking, and also predict academic achievement positively (Pekrun et al., 2002). In contrast, negative emotions are associated with task-irrelevant thinking, decreased interest, motivation and effort, and also predicted academic achievement negatively (Pekrun et al., 2002); specifically anxiety may hamper opportunities for learning by reducing motivation and engagement (Lehman et al., 2008). Moreover, Muis et al. (2015) found that positive epistemic emotions predict deep learning processes, like elaboration of the material and critical thinking, while negative epistemic emotions predict shallow learning processes, like sole rehearsal.

However, the relations between emotions and the quality of learning processes are not always entirely clear-cut. For example, individual negative emotions (here: anxiety) can predict deep processes (Muis et al., 2015); and there is also evidence that positive emotions (here: enjoyment) can predict learning outcomes negatively, and negative emotions (here: anxiety) can predict learning outcomes positively (Trevors et al., 2017). Possibly, positive emotions may widen learners' attention also towards not so central aspects of the learning tasks, and negative emotions may turn learners' focus towards the most salient information to be learnt, and depending on the learning context and the way that learning outcomes are measured, the unexpected results of emotions can be explained (Trevors et al., 2017). In another line of thought, it has been argued that also seemingly negative interactions or behaviors – such as physical disagreement – can be helpful for learners' collaboration (Fleck et al., 2009). In this sense, it may be assumed that, likewise, the expression of negative emotions possibly may have a productive function in group learning. Lastly, one can question if all 'undesired' outcome variables are actually hindering learning; for example, according to Greene (2015), both deep and shallow approaches are essential for mathematics learning.

Besides emotions on the side of the individual learner, emotions are also central for group learning processes. Learners experience social emotions, for instance, when comparing each other's performances, or evaluating the classroom atmosphere and the interactions with their peers (Järvenoja & Järvelä, 2005). Collaboratively constructing knowledge may be emotionally challenging at times (Järvenoja & Järvelä, 2009), requiring learners to regulate their own, their learning partners', and the whole group's shared emotions (Järvelä & Hadwin, 2013; Järvenoja & Järvelä, 2009).

While emotions can be thought of as inner and subjective phenomena, they usually play out through facial expressions and actions, or readiness for actions, such as approaching or retracting (Frijda, 1988). Thus, observable emotions (e.g., in behavior or facial expressions) are not only representing an internally perceived emotion, but expressing the emotion is a crucial aspect of the emotional experience itself (Niedenthal et al., 2005; Sakr et al., 2016). The body plays a central role in how humans deal with and process emotional information, making emotional processing embodied (Niedenthal et al., 2005). The extent, however, to which learners in a collaborative setting overtly express or control their emotions, differs between individual learners (Järvenoja & Järvelä, 2005). Moreover, Ekman (1992) reports that the different emotions classified as positive are much more difficult to distinguish from another (e.g., through distinct facial expressions) than the ones classified as negative, which may be explained by the evolutionarily higher importance of expressing and identifying negative emotions, such as anger, compared to positive emotions (Ekman, 1992).

In conclusion, the relations between different emotional processes and learning variables are complex (e.g., Muis et al., 2015; Trevors et al., 2017), but it can be noted that emotions play a crucial part for learners' experiences in a learning setting. Therefore, an effort should be made to involve emotions in the analyses of learning processes. Especially, in a co-present and embodied CSCL scenario, as in the case of the studies in this thesis, learners may express their emotions through their body, which would make the perceived and expressed emotions an integral part of the learning experience for the collaborative learners. Bodily expressed emotions have been included as a non-verbal process variable in studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4). The way of conceptualizing positive and negative emotions and representing these variables in coding schemes in this thesis can be found in section 6.2.1.

4.2.2. The relevance of task focus

What is task focus in the context of learning? *Task focus*, and the related terms *engagement* (Deater-Deckard et al., 2013) or *participation* (Weinberger & Fischer, 2006), refer to learners' efforts to solve the tasks in a learning environment (Baker & Lund, 1997; Deater-Deckard et al., 2013; Weinberger & Fischer, 2006). Learners' task focus may become evident through their verbal contributions, for example, learners' discourse may be on-task or off-task (Weinberger & Fischer, 2006). But the degree of task focus may also show through aspects of cognitive, behavioral, and affective engagement (Deater-Deckard et al., 2013). Learners may purposefully direct their attention towards the task (cognitive engagement), demonstrate fast and accurate motor responses for solving a task (behavioral engagement), and approach a learning environment with excitement and happiness (affective engagement) (Deater-Deckard et al., 2013). In contrast, learners may at times or continuously be off-task – the opposite of being task focused. Off-task behavior can be interpreted in terms of frustration because the task demands cannot be met (Webb, 1982), and this off-task behavior may also be frequent in computer-supported learning scenarios (Järvenoja & Järvelä, 2005). Encountering adequately difficult tasks, i.e. challenging but not overwhelming tasks, is assumed to increase learners' task focus (Deater-Deckard et al., 2013).

Task focus is a learning-relevant variable: Turning one's focus towards the tasks in a learning environment is a basic necessity for any learning to happen (Weinberger & Fischer, 2006). For example, prior research identified a positive relationship between observed engagement and test performance (Deater-Deckard, El Mallah, Chang, Evans, & Norton, 2014). Engagement in problem-solving, for example, multiple attempts for finding solutions, was found to contribute to learning gains for students with low prior knowledge (although not for students with high prior knowledge) (Hulse et al., 2019). In contrast, off-task behavior is negatively related to learning outcomes (Webb, 1991). Finally, in collaborative learning, task focused behavior is considered to be one of the important factors contributing to productive collaborative interactions (Baker & Lund, 1997).

In conclusion, successful learning processes are dependent on learners' ability to focus on the given tasks (e.g., Weinberger & Fischer, 2006). If learners do not want to or are not able to continuously engage with the learning environment, successful learning is unlikely. Different (variants of) learning environments may contribute differently to learners' task focus, and therefore, also task focused behavior, respectively off-task behavior, was included as one quality of the learning processes in co-present learning in studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4). The way of conceptualizing task focus as (lack of) off-task behavior and representing this variable in a coding scheme in this thesis can be found in section 6.2.2.

4.3. Supporting Learning Processes

The past sections outlined process variables which are relevant for shaping (collaborative) learning. These variables do not only span the epistemic quality of learners' verbal contributions or the extent that learners refer to each other (transactivity), but also non-verbal processes, such as task focused behavior or bodily expression of emotions. Successful collaboration cannot be taken for granted, as it requires learners to demonstrate sustained effort to relate to each other, and to coordinate their verbal exchanges and bodily activities with the goal of constructing knowledge together (Roschelle & Teasley, 1995). In the case of collaborative learners' talk, beneficial processes like presenting one's own understanding, asking and answering questions, giving explanations, resolving conflict, sharing and jointly evaluating understanding, elaborating scientific concepts, drawing conclusions, or taking shared decisions, may not spontaneously occur if learners are unsupported (e.g., Mercer, 1996; Falloon & Khoo, 2014; van Boxtel et al., 2000). For instance, in the study by Falloon and Khoo (2014), five year old learners collaboratively using tablet apps were mostly engaged in cumulative talk, followed by exploratory and disputational talk. In addition, in a classroom, children may have different ideas about the functions of discussions, for example, children may utilize discussions as a way for refining their understanding and clarify open questions, or they may rather want to proof their knowledge to their peers (Mercer, 1996).

As a consequence, a lot of research has been dedicated towards setting up beneficial conditions for learning, or supporting learners in engaging in productive learning processes (e.g., the ones presented in sections 4.1 and 4.2 above), often with a focus on fostering productive talk (e.g., King, 1990; Webb, 1991). Enhancing learning processes may be realized through the instructional or learning environment design, or through a teacher (Falloon & Khoo, 2014; Gijlers et al., 2013; Mercer, 1996; Tsovaltzi et al., 2017), for example, by eliciting elaborated explanations (e.g., Webb, 1991). The design of a learning environment, for example a tablet, can influence talk in the way that an open design, allowing for learners' own free input, increases the chance of productive collaborative, explorative, talk (Fallon & Khoo, 2014). Duschl (2008) argues for designing learning environments in a way that they facilitate "making students' thinking visible" (Duschl, 2008, p. 287) in order to enhance their epistemic discourse abilities. Mercer (1996) summarizes three main conditions that need to be met for fostering productive talk: First, talking is an explicit part of the learning activity and not just a by-product, second, the learning scenario facilitates cooperation – not competition, and third, the peers have a shared understanding of the task (Mercer, 1996).

The learning environment design may also be adapted by including scripting or scaffolding elements. Scripting approaches make explicit what is expected from collaborative learners and how they should interact (Dillenbourg, 2002), and are designed to support learners in engaging in the sort of activities and interactions which are conducive to learning (Jeong & Hmelo-Silver, 2016). Scripts can, for example, define the sequence of learning activities, or distribute roles and tasks (Dillenbourg, 2002; Weinberger, Stegmann, Fischer, & Mandl, 2007). Scaffolding elements are designed to guide learners while solving the task, and to support knowledge acquisition processes (Reiser, 2004). Scaffolding can, for example, simplify or structure a task, or emphasize critical aspects of a task by eliciting additional explanations from the learners (Reiser, 2004). Both scripting and scaffolding are means to support learners to engage in meaningful processes, and in practice, they can be realized through prompts (e.g., Reiser, 2004; Weinberger et al., 2005), such as short instructions or questions for the learners.

For instance, it has been found that in an individual learning scenario, question prompts targeting learning processes of observing, recalling, and concluding significantly improved learners' domain knowledge and performance in far transfer tasks, i.e., their potential for knowledge transfer (Demetriadis, Papadopoulos, Stamelos, & Fischer, 2008). In collaborative scenarios, learners may be prompted to ask questions, listen to their learning partner, and critically reflect on their contributions, in other words, the discussion is guided towards explorative talk (Fallon & Khoo, 2014), respectively towards higher levels of transactivity. Script prompts, sequencing different central aspects of a discussion, showed to successfully enhance learners' transactivity, especially when combined with worked examples (Vogel et al., 2013). Also productive epistemic activities may be fostered: For example, Weinberger et al. (2005) implemented epistemic prompts – adapted from Brooks and Dansereau (1983; as cited in Weinberger et al., 2005) – which prompted learners to include aspects of *theory*, *consequences*, *empirical findings*, and *individual judgment* into a jointly edited text document. In another study, an argumentation script for categorizing the arguments a learner writes improved learners' level of elaboration (Tsovaltzi et al., 2017). In a study with children, domain-related prompts improved participants' epistemic quality in dialogue (defining concepts), but did not affect the epistemic quality in their learning activity (creating a drawing including scientific concepts) (Gijlers et al., 2013). In addition to prompting for high-quality verbalizations, prompts can also target strategic behavior, for instance, by suggesting a certain sequence of activities or by structuring, respectively restricting, interactions with a learning environment (e.g., Bodemer, Ploetzner, Bruchmüller, & Häcker, 2005; Jackson, Krajcik, & Soloway, 1998; Sharples et al., 2015). To conclude, supporting learners with prompts often showed positive effects in prior research. However, this approach does not always work as intended and may have zero or even detrimental effects, for example, when it fails to elicit high-quality elaborations (Weinberger et al., 2005), or when naturally beneficial collaborative interactions are impeded through *over-scripting* (Dillenbourg, 2002).

In study III (Schmitt & Weinberger, 2019; section 7.3) of the present thesis, two sorts of prompts were developed and employed in the *Proportion* app, to investigate if and how added *structure* and a focus on *verbalizations* may affect non-verbal and verbal process qualities and learning gains, and possibly complement children's collaborative and bodily activities in an embodied learning environment. One prompt type aimed at fostering strategic behavior while working with the app, and the other prompt type aimed at fostering high quality verbalizations. These prompts were designed to create an environment in which learners feel engaged and positive towards their learning activities, and have explicit opportunities for expressing one's understanding of the tasks and underlying concepts, in order to promote knowledge construction (see also, section 5.4).

4.4. The Role of Process Heterogeneity in Collaborative Learning

Relevant aspects of learning processes – like learners' verbalizations, task focus, or emotions – do not only operate on the level of the individual learner. In a collaborative scenario, these processes play out for each individual learner, but they are also expressed and perceived in the context of the collaborative group, which may display more or less heterogeneous processes. Thus, the question arises, how these potentially diverging learning processes within a collaborative group may affect the collaborative processes and the success of learning.

In research on shared touch interfaces, equal / equitable access and participation seemed often to be considered as one of the key advantages of using such a technology, and as an indicator for good collaborative interactions (e.g., Jakobsen & Hornbæk, 2016; Jamil et al., 2017; Rick, Kopp, Schmitt, & Weinberger, 2015). In this line of thought, homogeneously distributed bodily activities between learners may be conducive to learning. In contrast, there is a whole body of research that points at inconsistent effects of learner heterogeneity in groups (e.g., Cheng, Lam, & Chan, 2008), or supports the assumption that heterogeneity can be advantageous (Webb, 1991), for example by fostering productive cognitive conflicts (Jorczak, 2011), or by enabling learners to complement each other's knowledge (Erkens, Bodemer, & Hoppe, 2016). However, this research on learner heterogeneity mostly focused on learners' *prior characteristics*, like different levels of ability. Little is known about heterogeneity in ongoing learning *processes*. For instance, learners in a group may display heterogeneous levels of bodily expressed emotions, or they may be more or less homogeneous regarding the epistemic quality of their verbalizations. Should a group rather be homogeneous or heterogeneous with respect to their learning processes?

To further our understanding of the central role of ongoing learning processes as well as questions of heterogeneity vs. homogeneity within collaborative groups for successful knowledge construction, this thesis also investigates the role of learning process heterogeneity. Specifically, study IV (Schmitt et al., in preparation; section 7.4) will address the effects of learning process heterogeneity on performance and cognitive processes variables in embodied collaborative learning.

4.5. Conclusion

This past chapter can be summarized by a quote of Dillenbourg et al. (2009, p. 6): "Collaboration per se does not produce learning outcomes; its results depend upon the extent to which groups actually engage in productive interactions". Productive interactions are characterized by:

- **High epistemic quality**: Verbal contributions which elaborate on and connect concepts (Tsovaltzi et al., 2017), and relate prior knowledge with learning tasks und underlying theories (Weinberger & Fischer, 2006)
- **High transactivity**: Learners explicate their knowledge in their verbal contributions, relate their contributions with each other and build up on each other, and engage in cognitive conflict by challenging the views of their learning partner (e.g., Gweon et al., 2013; Mercer, 1996; Teasley, 1997)
- **Sustained task focus**: Learners make a continuous effort to engage with the learning tasks, through on-task talk, or cognitive, behavioral, and affective task engagement (Baker & Lund, 1997; Deater-Deckard et al., 2013; Weinberger & Fischer, 2006)

Further, learners' emotional processes are central to learning, both on the individual level as well as in a collaborative setting. While there is evidence that positive emotions are positively related to desirable learning processes and outcomes (and negative emotions are related negatively) (e.g., Pekrun et al., 2002), opposite effects have also been found (e.g., Trevors et al., 2017), underlining the point of view that the relationship between these variables is complex and may depend on the specific learning context to a great extent.

Learning processes may benefit from additional prompts that support learners in their interactions (e.g., King, 1990), for example, to explain concepts to their learning partner. It remains unclear if prompting also has advantageous effects if applied in an embodied learning environment which is strongly based on bodily interactions. The effects of two sorts of prompts will be addressed in study III (Schmitt & Weinberger, 2019; section 7.3). Moreover, the expression or quality of different learning processes may be rather homogeneous or heterogeneous within a collaborative group; study IV (Schmitt et al., in preparation; section 7.4) will investigate this largely unexplored topic of learning process heterogeneity.

Regarding quantitative and qualitative measurement methods to capture learners' experiences and ongoing learning processes on multiple dimensions, chapter 6 will give an overview on related research and methods utilized by other researchers, and introduce a survey instrument and qualitative coding schemes which were developed and employed in the course of the present thesis.

Before elaborating on these methods, the following chapter 5 will give an overview on the learning domain and environment of this thesis: Proportional reasoning and the *Proportion* tablet app.

5. Proportional Reasoning and the Proportion Learning Environment

Studies II, III, and IV (see chapter 7) as part of this thesis utilize the *Proportion* learning environment. *Proportion* is an iPad application (app) originally designed and programmed by Dr. Jochen Rick at the Educational Technology department at Saarland University (e.g., Rick, 2012a), and further enriched as part of the research in the present thesis (see section 5.4). The following sections will introduce the domain of proportional reasoning (section 5.1), the idea and design of the *Proportion* learning environment (section 5.2), the learning setting and context of *Proportion* (section 5.3), and the experimental variation of *Proportion* with two sorts of additional prompts (section 5.4).

5.1. The Domain of Proportional Reasoning

This section will introduce the concepts around the domain proportional reasoning, along with related challenges and approaches to support its development.

The learning domain of the *Proportion* app is proportional reasoning, an important skill in the domain of mathematics (de la Torre, Tjoe, Rhoads, & Lam, 2013), but also everyday activities such as baking (Boyer et al., 2008). To think proportionally involves the ability to reason about, and possibly also manipulate, fractions (Rick et al., 2012), ratios (Jitendra, Star, Rodriguez, Lindell, & Someki, 2011; Tourniaire & Pulos, 1985), rates, and percentages (Jitendra, et al., 2011), i.e. dealing with multiplicative relationships (Boyer et al., 2008).

Proportional reasoning can be challenging for children (Boyer et al., 2008; Fujimura, 2001; Mix, Levine, & Huttenlocher, 1999; Rick, 2012a), as previously acquired skills conflict with thinking proportionally (Roschelle et al., 2010), for instance, counting and comparing absolute numerals or units rather than taking into account their proportional relation (Boyer et al., 2008).

In the primary school context of Saarland (the federal state in Germany in which this thesis was done), there are not many opportunities for the pupils to engage with concepts of proportions, as proportional reasoning is covered only peripherally in the primary school curriculum (Saarland Ministerium für Bildung, Familie, Frauen und Kultur, 2009). For instance, in grade three and grade four, one competency includes, among others, to “solve easy word problems regarding proportionality” (translated, Saarland Ministerium für Bildung, Familie, Frauen und Kultur, 2009, p. 19 and p. 27). Further, in third grade, students are supposed to know the “everyday fractions $\frac{1}{2}$, $\frac{1}{4}$, $\frac{3}{4}$ for lengths, weights and time spans” (translated, p. 20), for example that 500g equals $\frac{1}{2}$ kg (Saarland Ministerium für Bildung, Familie, Frauen und Kultur, 2009).

As there is evidence that proportional reasoning is present already in children as young as seven years old, at least in some intuitive form (Boyer et al., 2008), researchers have supported the idea of learning environments that build on this intuitive knowledge, for example, focusing on visual comparisons (Boyer et al., 2008), or providing hands-on experiences, for instance, with manipulatives (Tourniaire & Pulos, 1985). Learning environments building on an embodied understanding and involving bodily movements to physically experience the proportional relations include, for instance, learning with handheld devices (Mathematical Image Trainer, Reinholz et al., 2010, see also section 3.4), or magnets as manipulatives (Fujimura, 2001), and may enable "action-before-concept" learning (Reinholz et al., 2010, p. 1490). Also the *Proportion* app, which will be discussed in the following sections, draws on the idea of involving learners with their bodies to construct an understanding of proportions, and foster learners' proportional reasoning.

5.2. Idea and Design of the Proportion App

This section will introduce the general idea and design of the *Proportion* app based on Rick (2012a), Rick (2012b), and Rick et al. (2012).

In *Proportion*, two vertical bars (one orange, one blue) are presented next to each other. Each bar is associated with a numeral (see Figure 2). The main task is for the learner to manipulate the two bars, via touch / the fingers, i.e. to resize them until they represent their correct proportional relation as indicated by the numerals. As the task is about proportionality, and not absolute values, many different absolute bar sizes / bar configurations may solve the task, as long as the proportional relation between the bars is correct. According to ter Vrugte et al. (2015), the tasks of *Proportion* may be classified as *comparison problems*.

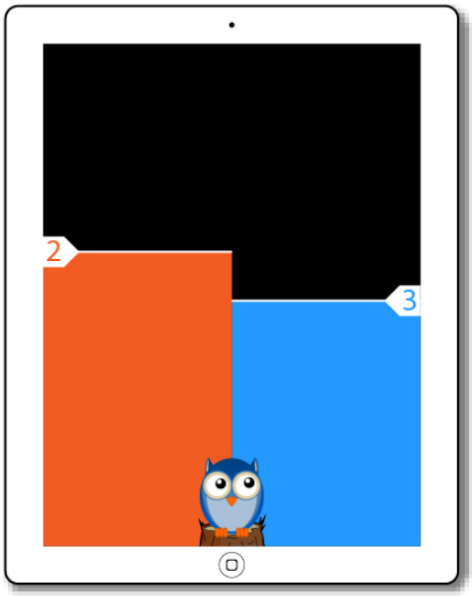


Figure 2. The basic *Proportion* interface. In this task, the two bars need to be adjusted to be in a proportional relation of 2 : 3.

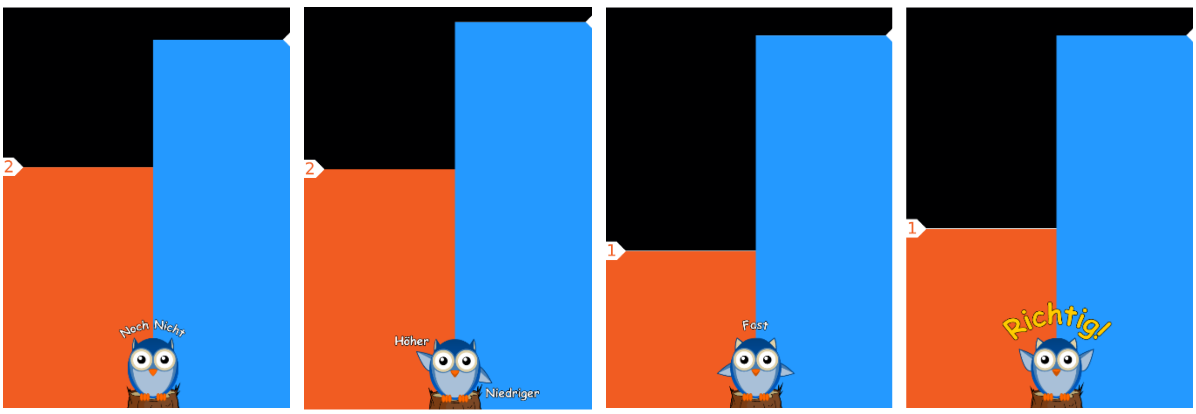
In the app version used for study II (Schmitt et al., 2019; section 7.2), *Proportion* is organized into 18 levels with 5 – 23 tasks each, with in total 168 tasks. In the app version used for studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4), new levels were added at the end, resulting in 21 levels with 5 – 23 tasks each, with in total 216 tasks. The tasks are becoming increasingly harder as learners progress through the levels, for instance, example tasks of level 1 are to configure the relations 1 : 2, or 4 : 3, example problems of level 21 are to configure the relations $\frac{2}{3} : \frac{1}{6}$, or $\frac{12}{5} : \frac{13}{7}$. Table 1 presents the levels with example problems and summarizes the main new concepts and challenges that are being introduced subsequently. Different forms of grids (no grid, fixed-position grid, relative lines, labeled lines) are laid over the basic interface in the different levels and were designed to facilitate varied strategies to solve the *Proportion* tasks (Rick, 2012a).

Table 1
Overview on Proportion Levels, Types of Problems, and New Challenges

App version in study II	App version in studies III and IV		
Level		Types of problems / new challenges	Example problems
1 – 4		Comparing two integers (min. number = 1, max. number = 5)	1 : 2; 3 : 4
5		Comparing two integers (min. number = 1, max. number = 9)	8 : 3; 1 : 7
6		Comparing two integers (min. number = 2, max. number = 40); New challenge: Representing numbers that seem too large to fit the screen	16 : 8; 15 : 5
7 – 9		Comparing two integers (min. number = 1, max. number = 300); New challenge: Dealing with large numbers	300 : 100; 10 : 20
10 – 12		Comparing two integers (min. number = 3, max. number = 23); New challenge: Comparing relatively large numbers which are not multiples of each other	14 : 11; 15 : 23
13 – 14		Comparing two fractions with the same denominator; New challenge: Encountering fractions (denominators can be ignored here)	$\frac{2}{5} : \frac{4}{5}; \frac{7}{11} : \frac{1}{11}$
15		Comparing two fractions with different denominators; New challenge: Understanding the role of denominators (a larger denominator does not make the value of the fraction larger)	$\frac{1}{5} : \frac{1}{2}; \frac{1}{3} : \frac{1}{4}$

Level: App version in study II	Level: App version in studies III and IV	Types of problems / new challenges	Example problems
16		Comparing integers with fractions, or comparing two fractions; New challenge: The combination of integers and fractions; integers could be thought of as fractions ($\frac{1}{1} = 1$)	$1 : \frac{1}{4}; \frac{1}{1} : \frac{1}{2}$
17 – 18	17 – 21	Comparing integers with fractions, or comparing two fractions; New challenge: Increasingly complex combinations which are not multiples of each other	$\frac{7}{2} : 3; \frac{1}{7} : \frac{2}{13}$

Proportion includes a small owl providing minimal feedback (see Figure 3): If the solution has not been configured yet, but the owl is touched, it displays 'not yet'. If the learners cannot solve the task after one minute, it helps by displaying 'higher' or 'lower' and pointing its wings up respectively down to indicate in which direction the bars should be moved. It displays 'close' when the configuration is almost correct, and it displays 'correct' and hoots cheerfully once the task is solved.



(a) (b) (c) (d)

Figure 3. The owl as pedagogical agent displaying 'not yet' (a), 'higher' / 'lower' (b), 'close' (c), and 'correct' (d).

Overall, the *Proportion* app can be characterized by its child-friendly design with bright colors and an animal avatar (pedagogical agent), and by its affordance for active experiencing through the body, embodied learning, and an intuitive grasp of proportional relations between quantities. Further, *Proportion* features individual game elements (tasks grouped into increasingly harder levels, rewarding feedback), without including many others (e.g., time running out) (Rick et al., 2015). These gaming features may be motivating for some students (Rick et al., 2015), without providing a fully gamified experience with its own strengths and weaknesses, which are beyond the scope of this thesis.

To give an idea to what extent the participants were able to solve the *Proportion* tasks, Table 2 provides an overview on descriptive statistics. It can be concluded that learners of all ability levels were able to use *Proportion* and made some progress at least, while no learner hit the ceiling and solved all of the problems.

Table 2
Descriptive Statistics of the Number of Proportion Tasks Solved by the Study Participants

Sample	Max. number of tasks in the app	M (SD) of tasks solved	Min. number of tasks solved	Max. number of tasks solved
Study II				
Dyads (*n* = 50)	168	108.5 (24.6) ≈ level 13	37 ≈ level 5	158 ≈ level 17
Individuals (*n* = 25)	168	89.2 (38.2) ≈ level 11	22 ≈ level 4	149 ≈ level 16
Study III dyads (valid *n* = 75)	216	99.5 (25.5) ≈ level 13	29 ≈ level 5	169 ≈ level 18

5.3. Proportion in Action: Learning Setting and Context

This section presents the learning setting for which *Proportion* was designed to be used and classifies the learning environment according to different frameworks.

Proportion is targeted at learners aged 9-10 years old (Rick et al., 2012), and specifically enables two co-present learners sharing the tablet in a face-to-face setting (Rick, 2012a), and collaboratively working with the tablet as “tiny tabletop” (Rick, 2012b, para. 1), which possibly affords the acquisition of shared embodied representations (see Barsalou, 1999). Multiple learners sharing the same multi-touch device also implies that the participants are physically close together, which may foster a high degree of awareness of each other's actions (Rick, 2012b).

The app runs on a multi-touch interface (the Apple iPad 2 tablet; "iPad 2 – Technical Specifications," 2013) affording direct and natural interactions of multiple fingers and / or multiple users (multi-touch). Figuring out the solutions requires active manipulation of the screen through the body, resizing the two bars with the fingers. The direct touch and active manipulation with the body is assumed to support learners in embodying proportional relations (Rick et al., 2012). An embodied understanding and feeling / intuition for the relations are becoming more important the more difficult the levels become, for example, when numbers are not multiples of each other anymore (see Table 1).

While learning with *Proportion*, a learner may (implicitly) form hypotheses and subsequently put them to the test. For example, a learner may have the hypothesis that $\frac{1}{3}$ (left bar) is larger than $\frac{1}{1}$ (right bar), because 3 > 1. Consequently, the learner sets the $\frac{1}{3}-$bar higher than the $\frac{1}{1}-$bar and checks if the solution is correct. It turns out to be not correct, but the learner may still keep the hypothesis and play around with the fine tuning of the bar sizes, while sticking to the fundamental assumption that the $\frac{1}{3}-$bar needs to be higher than the $\frac{1}{1}-$bar. After one minute of not reaching a correct solution, the owl will give hints, i.e., display "lower" on the left and "higher" on the right, and guide the learner to the correct solution. Once the learner realizes that a larger denominator means smaller – and not larger – quantity, the learner may form a new hypothesis based on that and try it out in the next tasks. Prior research showed that hypothesis testing is effective for fostering proportional reasoning (Asterhan, Schwarz, & Cohen-Eliyahu, 2013).

Proportion as a learning environment can be classified according to different frameworks:

Regarding its affordance for bodily activities, *Proportion* serves the function of online embodiment (Niedenthal et al., 2005; Wilson, 2002), respectively physical embodiment (Black et al., 2012), as the activity involves active manipulation through the body rather than mental simulation or imagination. Further, as the users themselves are directly involved with their own body, *Proportion* targets direct embodiment, and not surrogate or augmented embodiment, which would involve controlling a surrogate or an avatar (Black et al., 2012). In conclusion, according to the terminology of the Instructional Embodiment Framework (Black et al., 2012), *Proportion* could be classified as a direct physical embodied learning environment.

Regarding its group learning aspect, the *Proportion* learning scenario could be classified as co-present cooperative learning with defined and divergent tasks in the matrix of CSCL scenarios (Weinberger et al., 2018): While the *Proportion* tasks themselves are defined, multiple solutions (configurations of the bars, as well as verbal elaborations when prompted, see section 5.4) are possible and correct. Depending on the users, *Proportion* affords both cooperative as well as collaborative interactions: Learners may split the work and take turns in solving the problems (rather cooperative), or they may work with a high degree of interactivity, jointly manipulating the bars (rather collaborative) and discussing the underlying concepts, or they may engage in a flexible mix of both. Roschelle and Teasley (1995) write that phases where learners do not interact intensively with each other but explore things on their own are normal for collaborative interactions, as long as the learners come back together and share their new insights. Based on these considerations and the notion that the distinction between the concepts of collaborative vs. cooperative learning is not clear-cut and agreed on (see also, chapter 2), *Proportion* can be also considered a collaborative learning environment.

5.4. Additional Prompts in Proportion

In the course of the present thesis, extending the baseline *Proportion* interface as presented in section 5.2, two sorts of prompts were added and tested in study III (Schmitt & Weinberger, 2019; section 7.3): *Strategy prompts* and *verbalization prompts*. Prompting learners was considered to be required to make learning with the app more effective, and to support them in connecting the bodily experiences and app-related abilities to the abstract underlying mathematical principles.

Prompting learners may target different goals, for example, motivational aspects, domain-related support, support for coordination, and other target qualities. For example, initial ideas for prompting included motivational prompts (e.g., displaying ‘This was a very hard problem. Great that you solved it! Keep it up!’ when the solution took a long time), or specifically prompting very fast users (to make them reflect and not just “do”), or slow users (to support their understanding). In the course of this thesis, a final decision was taken for verbalization and strategy prompts. In short, the verbalization prompts request additional explanations from the learners, for example, ‘Explain to your learning partner: What did one need to do in order to solve the task?’ to make them abstract from the concrete task to the underlying principles. The strategy prompts support strategic behavior in the app, for instance, the strategy of first thinking, then manipulating (‘Tip for all tasks: First think and provide an estimate, then set the bars' correct height!’). More information on the pedagogical underpinning of the prompts can be found in study III (Schmitt & Weinberger, 2019; section 7.3; and Table 15). See Figure 4 for screenshots of the prompts in the *Proportion* app.

The prompts were designed to regularly appear after each first task of each level. As new levels often introduced new challenges (e.g., relating two integers, vs. relating two fractions with the same denominator, vs. relating two fractions with different denominators, see Table 1), it was considered useful to put on hold the bodily manipulations of the app after each first problem of each level, and confront the learners with questions and / or strategies in order to support them in handling the increasing difficulty. Prompting once per level should make sure, on one hand, that the learners have regular exposure to the prompts, and on the other hand, that the prompts will not appear too often with the danger of entirely changing the hands-on *Proportion* experience to a fully strategic learning environment relying on verbal discussions.

Each prompt type (verbalization vs. strategy prompt) included three different versions (see Table 15) which were iterated over the levels. This approach served two purposes: First, by not encountering the same prompt version multiple times one after another, it was aimed at keeping learners' motivation and interest towards the prompts high. Second, repeatedly encountering prompts over time was supposed to help learners to internalize and effectively use them.

In the case of the verbalization prompts, which requested an explanation by the learners, the owl was "listening" to the learners by pricking up its ears when the app was sensing sound (children's voices). However, the owl was not actually recoding or understanding what the children were saying. The owl would only let the children proceed after at least five seconds of time (otherwise it would say 'I have not understood this'), to make sure that the children are at least trying to give an explanation, instead of just moving on with the tasks. The verbalization prompts also indicated which participant should give the explanation by pointing its wing to the left or to the right (the directing was randomized for each prompt appearance).

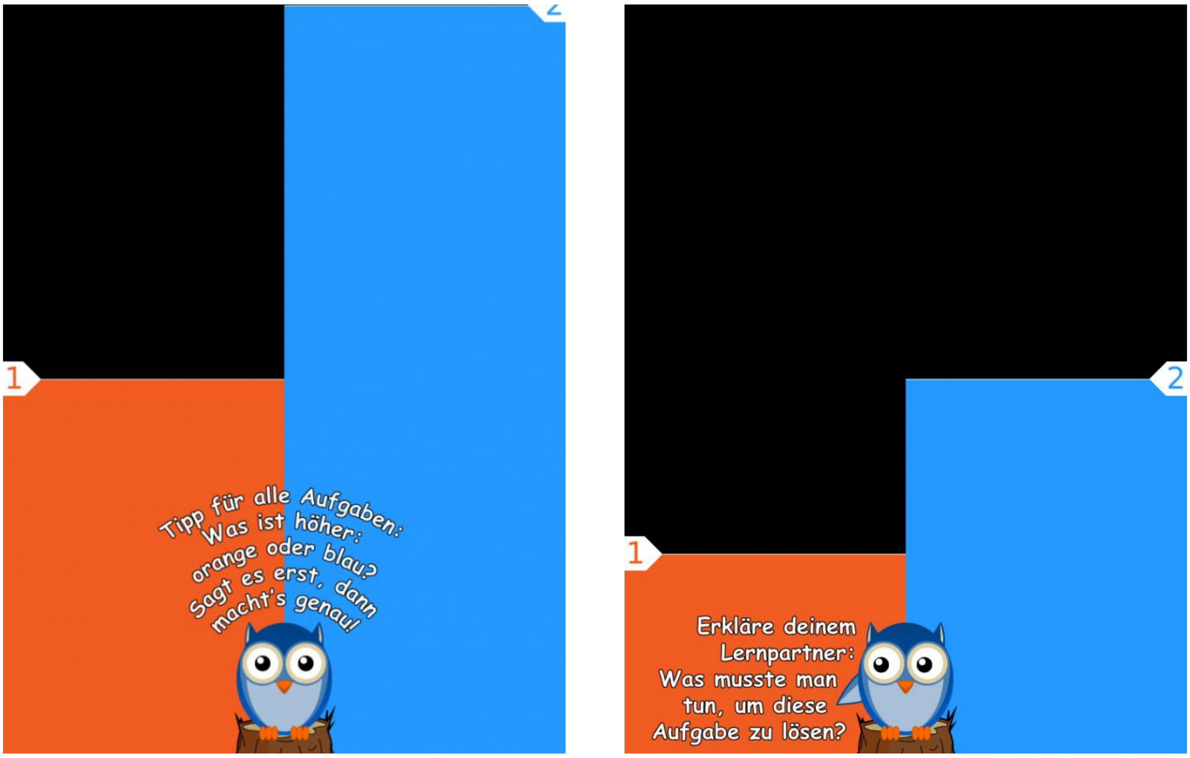


Figure 4. The *Proportion* interface with a strategy prompt (version A, left) and a verbalization prompt (version A, right).

The two prompt types were integrated into the app in a 2 × 2 design with the factors *verbalization prompts* (*VERB*; present / not present) and *strategy prompts* (*STRAT*; present / not present), resulting in four experimental conditions:

- *Control condition*: no prompts
- *VERB*: only verbalization prompts
- *STRAT*: only strategy prompts
- *VERB-STRAT*: combination condition: verbalization plus strategy prompts

In the combination condition, the verbalization prompt appeared first, followed by the strategy prompt. The programming of the prompts and their technical integration into the app was done by Dr. Jochen Rick. The experimental study on the effects of these prompts is published in Schmitt and Weinberger (2019) (study III; section 7.3).

5.5. Conclusion

Proportional reasoning is an important but challenging domain in mathematics (e.g., Mix et al., 1999). The development of proportional reasoning abilities may benefit from hands-on bodily experiences (e.g., Reinholz et al., 2010), for instance, through the *Proportion* app. The *Proportion* learning environment challenges young users with proportion tasks of increasing difficulty which can be solved by physically manipulating two bars on the touch screen. *Proportion* is a collaborative learning environment for two learners and targets direct physical embodiment. In addition to the baseline learning environment, it may make sense to provide further support in the form of prompting learners for strategic behavior and verbalizations. Additional prompting in *Proportion* is realized in study III (Schmitt & Weinberger, 2019; section 7.3).

Before presenting the four studies of the present thesis in chapter 7, the following chapter 6 will give an overview on research methods in the context of summative assessment of children's experiences, and qualitative process-oriented research with children.

6. Methods for Researching Children's Co-Present Learning with Technology

In the current chapter, two methodological aspects which are central to the research of this thesis are being discussed. As the studies done in the course of this thesis investigate children's learning, there is a need for research methods that are particularly suited for research with this young (around 10 years old) target group. Section 6.1 will introduce the challenge of sampling children's summative evaluations via questionnaires, motivating a need for child-friendly questionnaire design. In section 6.2, the development of coding schemes for doing qualitative video data analysis of the learning processes on multiple dimensions (i.e., different facets of verbal and non-verbal behavior) will be explained and justified. Developing and applying qualitative analysis techniques to fairly large amounts of data was one of the major challenges in the research process of this thesis.

6.1. Questionnaires for Summative Assessment

Closed item questionnaires are suitable for summative evaluation at the end of a learning intervention, and compared to other methods such as interviews or video observations, they are efficient for the research participants to respond to, as well as for the researchers to analyze the results. However, questionnaire responses may lead to a variety of response biases, for example, social desirability, acquiescence, satisficing, etc. (Moosbrugger & Kelava, 2008). While traditionally it was questioned if children are capable to respond to closed item questionnaires at all, more and more researchers are willing to include children as respondents, and consequently, develop appropriate instruments (e.g., Bell, 2007). Questionnaire instruments for children in the context of technology and learning are, for instance, *This or That* (Zaman, Abeele, & De Grooff, 2013), *Fun Sorter* and *Again Again* (Sim & Horton, 2012), and the *Smileyometer* (Read, 2008). While all of these instruments have noticeable strengths in addressing the challenges of surveying children, they also come with problems, for example, extremely positive responses with the Smileyometer (e.g., Johnson, Shum, Rogers, & Marquardt, 2016).

In conclusion, there is a need to further work on developing questionnaire instruments for young learners. As part of this research in this thesis, the BiCo (Bipolar Continuous Rating Scale) was developed and put to the test. Study II (Schmitt et al., 2019; section 7.2) contains a thorough review of the challenges stated above and the subsequent development of the BiCo instrument along with results on its utility.

6.2. Multimodal Data Analysis

In addition to summative evaluations at the end of an intervention, analyzing learning processes as they unfold during the learning phase provides valuable insights. Understanding learning processes in co-present embodied learning environments necessitates a multimodal analysis that goes beyond analyzing written text (as in computer-mediated scenarios), and needs to include other modalities as well, such as learners' gestures or speech (Gweon et al., 2013).

In the present thesis' research, multimodal video analysis was employed to analyze learning processes on multiple dimensions in the embodied learning environment *Proportion* (see chapter 5). In order to grasp the multifaceted interaction and learning processes in such an environment, a focus was laid on the bodily expression of emotions (section 6.2.1), bodily expressed focus on the task (section 6.2.2), and the quality of learners' verbal exchanges (transactivity and epistemic quality, section 6.2.3).

In the following sections, I will discuss different ways in which these variables have been conceptualized and measured in prior research and, subsequently, introduce the coding schemes which have been applied in the research of this thesis.[2]

6.2.1. Measuring bodily expressed emotions

For measuring emotions, a wide variety of methods has been utilized in prior research, each with their specific strengths and weaknesses. In a review about measurement methods for emotions, Mauss and Robinson (2009) include the following methods: Self-report (e.g., a questionnaire), autonomic measures (e.g., heart rate, sweat), startle response magnitude (e.g., eye blink), brain states (EEG or fMRI measures), and behavior (vocal, facial, and whole-body characteristics and behavior) (Mauss & Robinson, 2009). Deater-Deckard et al. (2013) name verbal and facial expressions as indicators for positive and negative emotions. Finally, emotions may also become evident through learners' collaborative discussions (Isohätälä, Näykki, Järvelä, & Baker, 2018), or by analyzing learners' think-aloud protocols (Muis et al., 2015).

Emotions are expressed through our bodies (Ekman, 1992; Frijda, 1988; Sakr et al., 2016), and bodily expressed emotions may be especially salient for young learners who may not feel a

[2] The coding schemes for bodily expressed emotions and task focus were developed by myself in collaboration with other students prior to this doctoral thesis work. They have been shortly presented and employed for the data analysis in my master's thesis (Schmitt, 2015). Here, I present these coding schemes in much more detail, review related approaches, and integrate them into a larger research context, in order to better understand the methodology and results in studies III and IV (sections 7.3 and 7.4). The coding schemes for transactivity and epistemic quality have been newly developed in the course of the present doctoral thesis.

high pressure to conceal the expression of their experienced emotions. Consequently, in studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4) of this thesis, the emotions variables are based on video-coding of observable bodily states and behavior during the learning activity. While an alternative method could have been to measure emotions through self-report (e.g., Deater-Deckard et al., 2013; Trevors et al., 2017), for example, through a questionnaire at the end of the learning activity, the qualitative approach of video-coding was chosen for two reasons: First, the focus of interest specifically was regarding the *bodily expression* of emotions; second, this approach allowed for sampling the emotional behaviors *during* the learning activity and not only retrospectively (see also, Sakr et al., 2016).

From the video observations, it became evident that the typical emotions the young participants show closely relate to their success or perceived challenges during their engagement with the learning environment. They show pride and happiness when they solve a task, which, for instance, materializes as clapping in the hands, throwing the hands up in the air, or showing thumbs-up. They also show anger when a task seems not solvable, for instance, through threatening gestures (threatening the tablet). Other than anger, they also show signs of frustration and hopelessness when the task demands become overwhelming and they cannot make progress, for instance, by showing their palms upwards, which seems to indicate that they do not know how to proceed. These observations are in line with, for example, Lehman et al. (2008), who report happiness, anxiety, confusion, and, to a smaller extent, frustration as the most commonly observed emotions during a learning situation.

The bodily expressed emotions were categorized into two larger clusters: *positive emotions* vs. *negative emotions*. This decision can be grounded into three considerations: First, also related research conceptualized emotions as largely falling into these two clusters of positive vs. negative emotions (e.g., Deater-Deckard et al., 2013; Frijda, 1988; see also section 4.2.1), and it has been shown that positive vs. negative emotions relate to learning outcomes differently (e.g., Pekrun et al., 2002; Trevors et al., 2017; see also section 4.2.1), making an analysis with this distinction relevant. Second, previous attempts to measure basic emotions during learning according to Ekman (1992) were less successful as these sort of emotions do not seem to occur to a high extent in a learning situation (Lehman et al., 2008). Third, qualitatively measuring emotions from video data and maintaining high objectivity and reliability is challenging. Consequently, also from a pragmatic point of view, a simple coding scheme that would allow for establishing sufficient inter-rater reliability was preferred.

Related research on observing emotions through behavior is sparse (Mauss & Robinson, 2009), and the description of methods is often not very specific. For example, Lehman et al. (2008) speak of “gross body movements” (p. 52) as indicators for emotions, but do not go into more detail which body movements they refer to. Similarly, Sim, MacFarlane, and Read (2006) speak of “positive body language” (p. 241) as one indicator for “enjoyment and engagement” (p. 241), but do not define what they consider to be an instance of positive body language. In a more recent publication, Sakr et al. (2016) see bodily moves as indicative for emotions, and give a couple of examples, for instance, clenching the fist as an indicator for joy or triumph, or jumping up and down as an indicator for excitement.

Building on a body of research on analyzing emotional processes, there is still a need to further specify the behavioral indicators for emotions. So far, these indicators were often not defined in detail, and therefore, these methods are not readily applicable for conducting own research. Furthermore, the variables of interest are partly conceptualized very wide. For instance, Sim et al. (2006) focused on “indicators of enjoyment and engagement, such as comments, smiles, laughter, or positive body language, and also signs of lack of enjoyment and frustration, including, for example, sighs, and looking around the room” (p. 241). This way of analysis does not distinguish between the aspects of emotions (“enjoyment”, “lack of enjoyment”, “frustration”) and focus on the task (“engagement”), and also the selected indicators are a mix of bodily behavior (“positive body language”, “looking around the room”), facial expressions (“smiles”), and oral expressions (“comments”, “laughter”, “sighs”) (p. 241). Similarly, in the context of the research by Deater-Deckard et al. (2013), next to *cognitive* and *behavioral engagement*, one aspect of engagement (or *task focus* according to the terminology used in this thesis) is *emotional engagement* with positive and negative emotions as indicators of engagement. Such an approach of a more holistic qualitative description may make sense depending on the context of the research, for example, for a richer view of a small data set (e.g., n = 25 in Sim et al., 2006). However, in the research done in this thesis, the variables were more strictly distinguished, i.e., positive emotions, vs. negative emotions, vs. task focus, and separately analyzed using individual coding schemes, in order to achieve high objectivity and reliability, and to make them applicable to a large data set (n = 162 in Schmitt & Weinberger, 2019 / study III, section 7.3).

In conclusion, the cited related work could serve as a starting point and inspiration, but needed to be expanded and specified in an own coding scheme. Thus, the resulting coding scheme for measuring the bodily expression of emotions (see Table 3) was partly an expansion of prior research, and partly inspired by the *Proportion* video data and the behaviors that were observable there. As the emotions coding scheme was workable in the data analysis in Schmitt (2015), it was adopted without changes for the analysis in the present doctoral research.

The emotions coding scheme (see Table 3) contains the chosen indicators for positive and negative emotions, as well as bodily states and behaviors that we frequently encountered but explicitly not conceptualized as indicative for a positive or negative emotion (*not counted*). For concrete examples of bodily expressed emotions found in the studies with *Proportion*, see Figure 5.

Table 3
Coding Scheme Developed for Measuring Emotions, as Published in Schmitt and Weinberger (2019) / Study III

Variable	Indicators	Not counted
Positive emotions	- clapping into one's own or the learning partner's hands - throwing hands up in the air - clenching the fist - showing thumbs-up	- change of body position - soft / indifferent hand movements - scratching the forehead - folding arms - shaking the head - knocking on the iPad or the table - letting the hand(s) fall down - showing the index finger - smiling / laughing
Negative emotions	- threatening the iPad - facing the palms upwards - dismissive hand gesture - face-palming	

For the research done in the course of this thesis, specifically for studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4), the coding scheme for emotions had to be trained to be operable by multiple coders (one student assistant and I), in order to reach sufficient inter-rater reliability values (Krippendorff's alpha; Krippendorff, 2012). The alpha values, assessed for 10 % of the analyzed data, were good with $\alpha = .81$ for positive emotions and $\alpha = .85$ for negative emotions (Schmitt & Weinberger, 2019 / study III, section 7.3).

(a) (b)

Figure 5. Examples for the bodily expression of emotions from the *Proportion* video data. In image (a), the participant on the left throws her hands up in the air, an indicator for the expression of a positive emotion. In image (b), the participant on the left faces his palm upwards, an indicator for the expression of a negative emotion.

6.2.2. Measuring task focus

In addition to bodily expressed emotions, another variable of learners' processes targeted their *task focus*. Learners' task focus can be measured by analyzing, for instance, learners' discourse, which may be on-task or off-task (Weinberger & Fischer, 2006), but also learners' behavior (Deater-Deckard et al., 2013). In the most basic level of task focus, one may measure if at all and how often a learner starts to engage with a learning environment, for example, logs into the system (Weinberger & Fischer, 2006). Next, once the learner has entered a learning environment, examples for behavioral indicators of engagement are the motor reaction time (Deater-Deckard et al., 2013), or visual focus / eye contact (Liu et al., 2009). Cognitive engagement can be assessed by means of learners' self-report of perceived attention, and affective engagement can be assessed by observing learners' facial expressions, or by learners' self-report of perceived positive and negative emotions (Deater-Deckard et al., 2013).

One coding scheme for measuring learners' engagement is *eSOCS* (*Engagement States Observational Coding Scheme*; Deater-Deckard et al., 2011). *eSOCS* comprises 13 aspects of engagement that the observer can rate on a scale from one to seven. The engagement indicators as proposed by Deater-Deckard et al. (2011) are summarized in Table 4.

Table 4
Summary of Aspects and Indicators of Learners' Engagement from the eSOCS Coding Scheme by Deater-Deckard et al. (2011)

Aspects of engagement in *eSOCS*	Indicators for engagement in *eSOCS*
Distraction / attention	The observer rates the proportion of distracted vs. attentive behavior, i.e. off-task vs. on-task, on a scale from 1 (constant distraction) to 7 (constant attention)
Positive affect	The observer rates the frequency of instances of positive emotions, for example facial expressions or excitement, on a scale from 1 (no instances) to 7 (instances constantly occurring)
Touching of device and / or task materials	The observer rates the frequency of touching the learning materials, for instance a device, on a scale from 1 (no touches) to 7 (constant touches)
Persistence	The observer rates how persistently the learner engages with the tasks, i.e. without breaks, on a scale from 1 (the learner does not engage with the task at all) to 7 (the learner shows continuous persistence)
Anger / frustration [1]	The observer rates the frequency of instances of these negative emotions[1], i.e. facial or verbal expressions, on a scale from 1 (no instances) to 7 (instances constantly occurring)
Gross motor movement	The observer rates the frequency of minor body movements, for instance pointing, and major body movements, for instance jumping, on a scale from 1 (no instances) to 7 (instances constantly occurring)
Anxious / nervous [1]	The observer rates the frequency of instances of these negative emotions, i.e. facial or verbal expressions, on a scale from 1 (no instances) to 7 (instances constantly occurring)
Fine motor movement	The observer rates the frequency of fine motor movements regarding the learning materials, on a scale from 1 (no instances) to 7 (instances constantly occurring)
Aggression [2]	The observer rates the frequency of aggressive behavior and verbalizations on a scale from 1 (no instances) to 7 (instances constantly occurring)
Verbalizations during task [3]	The observer rates the frequency of on-task talk on a scale from 1 (no instances) to 7 (constant on-task verbalizations)
Intrusiveness [4]	The observer rates the frequency of intrusive behavior, for example interrupting or controlling the learning partner, on a scale from 1 (no instances) to 7 (instances constantly occurring)

Aspects of engagement in *eSOCS*	Indicators for engagement in *eSOCS*
Responsiveness to partner [4]	The observer rates the frequency of responsive behavior, for example reacting to the learning partners' contributions, on a scale from 1 (no instances) to 7 (instances constantly occurring)
Independence / Autonomy [4]	The observer rates the frequency of leading or dominant behavior on a scale from 1 (no instances) to 7 (instances constantly occurring)

Note. The aspects of engagement (left column) appear as they are named in eSOCS (Deater-Deckard et al., 2011).
[1] According to Deater-Deckard et al. (2013), anxiety as well as frustration can be both subsumed under negative emotions.
[2] The aggression can be directed towards the learning materials or the learning partner (Deater-Deckard et al., 2011).
[3] The verbalizations can be directed towards one-self or the learning partner (Deater-Deckard et al., 2011).
[4] Only applicable for learners working in dyads (Deater-Deckard et al., 2011).

In addition to the aspects presented in Table 4, which can be coded for each individual, Deater-Deckard et al. (2011) present more aspects of engagement that are coded on a group level: *Cooperation*, *competition*, *conflict*, and *reciprocity*. However, these aspects are not central to the concept of task focus as presented in section 4.2.2 and will not be further elaborated on in this thesis.

The coding schemes utilized in the present thesis can be related to the *eSOCs* coding scheme: The *eSOCS* categories of *anger / frustration*, *anxious / nervous*, and *aggression* (Deater-Deckard et al., 2011) are reflected and combined in the coding scheme for *negative emotions* (see Table 3), presented in the previous section 6.2.1. The *eSOCS* category of *responsiveness* (Deater-Deckard et al., 2011) is analogous to the conceptualization of *transactivity* in this thesis (see the upcoming section 6.2.3). Regarding the *task focus* variable as it is conceptualized in the context of this thesis, similarities can be drawn to Deater-Deckard et al.'s (2011) *eSOCS* categories of *distraction / attention*, respectively *persistence*. In studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4) of this thesis, task focus was measured through video-coding of observable line of vision during the learning activity, i.e., the line of vision served as a bodily indicator of the learners' focus on the task (see also, Chung et al., 2013). In that sense, the operationalization of task focus is at the intersection of cognitive and behavioral engagement as defined by Deater-Deckard et al. (2013), because the cognitive aspect (attention on the task) is measured by means of a bodily indicator (line of vision).

Similar to the *eSOCS* coding scheme (Deater-Deckard et al., 2011), which rates the proportion of attentive / on-task behavior and persistent engagement vs. distracted / off-task behavior with a lot of breaks, the coding scheme for measuring task focus used in this thesis conceptualizes task focus (on-task behavior) as active or passive task engagement: Active engagement includes to actively manipulate the tablet app with the own fingers, to verbally interact with the learning partner, or to ask a task relevant question to the experimenter (rare). Passive engagement includes to monitor the learning partner's manipulation of the tablet app, indicated by student gaze focused on the tablet, or to listen to instructions provided by the experimenter (rare). It was not distinguished between active and passive engagement with the task, both were considered on-task behavior. The very quick and fluent transitions between active and passive phases made a separate analysis impossible. Moreover, observations from the video data did not suggest that physically dominant behavior of one learning partner would have been a frequent phenomenon. In general, the children shared active access to the tablet very evenly. For a concrete example of an instance of off-task behavior found in the studies with *Proportion*, see Figure 6.

Regarding the task focus variable, off-task behavior (instead of on-task-behavior) was coded and counted for the reason that on-task behavior was most of the time the norm, and off-task behavior rather the observable deviation from the norm. This approach can be substantiated by a detailed analysis of one dyad's off- and on-task behavior during learning with *Proportion* in Schmitt (2015): It was found that the average length of an instance of an off-task episode lasted 1.79 seconds, with in total 84 seconds of off-task behavior, resulting in 3.5% off-task behavior time during the entire learning period of 40 minutes. Also the task focus coding scheme was workable in the data analysis in Schmitt (2015), and therefore, adopted without changes for the analysis in the present doctoral research.

Verbal interactions with the learning partner were always classified as on-task behavior, regardless of the content of that talk. This approach was taken because the focus in this particular coding scheme lies on the bodily indicators of task focus, and not on the content or quality of the verbalizations (which is covered in the coding schemes for epistemic quality and transactivity presented in section 6.2.3). The task focus coding scheme regards any verbal act as on-task, which does not automatically include purely on-task talk. For successful collaborative learning, not only on-task-talk is an important contribution, but also other talk (*small-talk*) to ease the social side of the interactions (Thornborrow, 2003). Therefore, from a purely behavioral perspective, any talk with the learning partner can be regarded as working towards the goal of accomplishing the collaborative task, therefore being on-task. The distinction between on- and off-task talk is an integral part of the epistemic quality dimension (see Table 8), which regards the purely cognitive level of the dialogue. Descriptively, it can be reported that off-task talk in the epistemic dimension is not very frequent, accounting only for 1.4 % of the utterances (Schmitt & Weinberger, 2019 / study III, section 7.3; or Table 16). In conclusion, verbally interacting with the learning partner (Table 5) is almost perfectly on-task talk.

Table 5
Coding Scheme Developed for Measuring Task Focus, as Published in Schmitt and Weinberger (2019) / Study III

Variable	On-task behavior	Off-task behavior
Task Focus	- active manipulation of the bars in the app - monitoring of the learning partner manipulating the bars while the eyes are focused on the iPad - verbally interacting with the learning partner - interactions with the experimenters (asking questions related to *Proportion*, or listening to and watching the experimenters in case they gave instructions)	- looking around the classroom - looking into the camera - interactions with participants outside the own group

Again, for the data analysis of the research in this thesis, specifically for studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4), the coding scheme for task focus had to be trained to be operable by multiple coders (one student assistant and I), in order to reach sufficient inter-rater reliability values (Krippendorff's alpha; Krippendorff, 2012). The alpha value for task focus, assessed for 10 % of the analyzed data, was good with α = .92 (Schmitt & Weinberger, 2019 / study III, section 7.3).

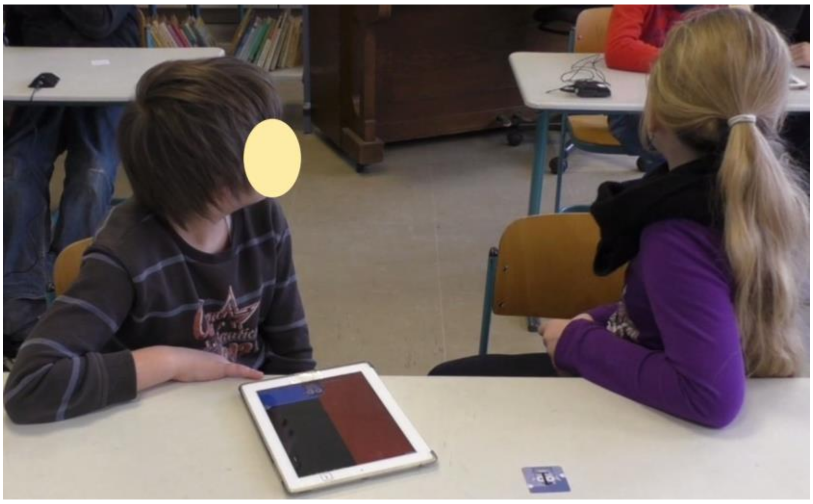


Figure 6. Example for an instance of off-task behavior from the *Proportion* video data. The participants look around in the classroom instead of engaging with the task.

6.2.3. Measuring the quality of dialogue

Analyzing learners' verbalizations in discussions, for example, their epistemic moves, argument structure, or social modes (transactivity), has a long tradition in CSCL research (e.g., Berkowitz & Gibbs, 1983; Teasley, 1997; Weinberger & Fischer, 2006). Research on the role of verbalizations can be experimental, i.e., identifying the experimental conditions that foster productive talk and learning outcomes, or observational, i.e., detailed *process* analyses of the nature of talk (Mercer, 1996). CSCL environments can be experimentally manipulated to facilitate certain types of talk, which then enables the researcher to examine the relationships between the differing emerging types of talk and learning outcomes (Baker & Lund, 1997). Moreover, the chosen approach can be rather qualitative, i.e., closely examining transcripts of learners' talk, or rather quantitative, for instance, counting the occurrence of certain key words (Mercer, 2004). Learners' talk is usually so complex that it requires a rich analysis from multiple angles, for example, with an argumentative and an epistemic focus, to understand it (Pontecorvo & Girardet, 1993). The following sections will give an overview how prior research has addressed this challenge and present the different perspectives on talk that researchers have been taking:

(a) **Disputational, cumulative, and explorative talk**: One approach for assessing quality of talk builds on Mercer's (1996) three categories of talk (disputational, cumulative, and explorative). While these categories were not meant to be used as a coding scheme directly (Mercer, 1996, 2004), they were utilized and further differentiated in later research. For instance, Falloon and Khoo (2014) adapted the categories towards a comprehensive coding scheme with subcodes and descriptions for each subcode. For example, explorative talk was further differentiated into *critically constructive*, *negotiated / debated*, and *justification sought / given* (Falloon & Khoo, 2014). This coding scheme then was applied to a learning scenario, where five-year-olds collaboratively used educational tablet apps.

(b) **Quality of questions and explanations**: King (1990) coded learner's talk based on talk categories presented by Webb (1989) into six categories in total: Giving vs. receiving *explanation* vs. *low-level elaboration* (resulting in four categories), and asking a question of the type *recall* vs. *critical thinking* (resulting in two categories). Basically, the scheme regarded different qualities within the moves of giving vs. receiving information and asking questions.

(c) **On- and off-task talk**: In another research approach, the focus was laid on aspects of children's on-task and off-task talk (Thornborrow, 2003). In this analysis, *on-task talk* is characterized by utterances that directly deal with solving the learning tasks in a goal-oriented manner; *off-task talk* is rather conversational and social in nature and less goal-oriented with respect to the task solution processes (Thornborrow, 2003). Also Baker and Lund (1997) distinguish off-task talk, which is social in nature, like joking, from on-task talk that is targeted at solving the learning tasks collaboratively. The on-task talk is then further differentiated into five sub-categories, for example, *interaction control* or *reflective evaluations* (Baker & Lund, 1997).

(d) **Content and relational aspects**: Janssen et al. (2011) coded learners' talk on a *content* vs. a *relational* dimension, and further include two dimensions regarding *coordination of content vs. relational aspects*, resulting in total in 19 different categories (this coding scheme was based on Baron & Kenny, 1986; as cited in Janssen et al., 2011). For example, the coordination of relational aspects included the categories *planning* (which strategies to use, how to work together), *monitoring* (the group processes and progress), and *positive and negative evaluations* (of the collaborative experience) (Janssen et al., 2011).

(e) **Purpose of the utterance**: Liu et al. (2009) coded learners' talk based on a coding scheme by Liu and Tsai (2006; as cited in Liu et al., 2009) into four categories: *Issues* (bringing up questions and possible strategies for task progress), *positions* (response to an issue, presentation of possible solutions), *arguments* (supporting or challenging a *position*), and *group development* (focus on the quality of collaboration and possible strategies to improve it) (Liu et al., 2009).

(f) **Transactivity**: Berkowitz and Gibbs (1983) produced a coding scheme for measuring the degree of transactivity with in total 18 sub-categories. The categories are considered to be of differing qualities (lower vs higher); examples for transactive moves of a relatively low quality are *feedback request* or *juxtaposition*, and examples for transactive moves of relatively high quality are *refinement* or *reasoning critique* (Berkowitz & Gibbs, 1983). Vogel et al. (2013) conceptualized transactive argumentation as "statements (…) that built on partner's contribution in an argumentative way" (p. 530) with sub-categories such as *criticizing* and *synthesizing*.

(g) **Epistemic moves**: Pontecorvo and Girardet (1993) use the following codes to assess epistemic operations: *Definition*, *categorization*, *predication*, *evaluation*, and *appeal to*, for example an analogy.

Weinberger and Fischer (2006) present an entire framework for analyzing learners' discussions in technology-supported scenarios. This framework is constituted of multiple dimensions: *Participation dimension*, *epistemic dimension*, *argument dimension*, and the *dimension of social modes of co-construction* (Weinberger & Fischer, 2006):

(a) **The participation dimension** captures to what extent the participants are invested in the task and contribute towards solving it, for instance through their verbalizations (Weinberger & Fischer, 2006). A second aspect of participation is the heterogeneity of participation within a group (Weinberger & Fischer, 2006). Heterogeneous learning processes are investigated in study IV (Schmitt et al., in preparation; section 7.4).

(b) **The epistemic dimension** reflects the content of learners' verbalizations and distinguishes between off-task and on-task talk, and subsequently, between different levels of quality of on-task talk (Weinberger & Fischer, 2006). The different levels of quality regard if learners' verbalizations are targeted on the concrete problem to be solved (*problem space*), the underlying theoretical assumptions (*conceptual space*), or if the verbalizations draw connections between the problem and related theories (*relations between conceptual and problem space*), or if the verbalizations reflect a connection between prior knowledge and the problem (Weinberger & Fischer, 2006).

(c) **The argument dimension** regards several aspects of a good argumentative structure, for example, if the learners provide claims with qualifiers and grounds, or if the learners pose counter-arguments or try to integrate different arguments (Weinberger & Fischer, 2006).

(d) **The dimension of social modes of co-construction**, also called transactivity, reflects to what extent the learners build up on each other, refer to each other, and critically incorporate the verbalizations of the learning partner (Weinberger & Fischer, 2006). Different levels of this dimension include pure *externalization* of thoughts, and higher-order moves like *integration* or *conflict-oriented consensus building*. These categories of transactivity are conceptualized as having a higher or lower quality, i.e., being more or less transactive (Noroozi et al., 2013; Weinberger & Fischer, 2006; see also, Gweon et al., 2013, and Weinberger et al., 2005).

The qualitative analysis in this thesis is majorly inspired by this framework of Weinberger and Fischer (2006): The participation dimension is reflected in the task focus variable (see previous section 6.2.2). Regarding the epistemic dimension and the transactivity dimension (*social modes of co-construction*), two coding schemes were developed in the course of the present thesis, which utilized and adapted the coding schemes presented in Weinberger and Fischer (2006). The argument dimension was not adopted as argumentation skills were not relevant to this thesis.

The coding scheme by Weinberger and Fischer (2006) has been applied (and adapted) in various research projects, many of them in the context of adults' computer-mediated text-based collaboration (e.g., Noroozi et al., 2013; Stegmann, Weinberger, & Fischer, 2007; Tsovaltzi et al., 2015; Tsovaltzi et al., 2017; Weinberger, Stegmann, & Fischer, 2010), often with a focus on only one or two dimensions of the original coding scheme. One major challenge for the development of epistemic quality and transactivity coding schemes for the present research was to translate this existing framework by Weinberger and Fischer (2006) into a version that would be suitable for assessing the dialogue of children who are collaborating in a co-present scenario. One such scenario of children learning in a co-present collaborative setting in prior research was the study by van Dijk, Gijlers, and Weinberger (2014). The application of the coding scheme by Weinberger and Fischer (2006) in this study by van Dijk et al. (2014) required a noticeable level of adaptation to fit to that specific context. For instance, the epistemic dimension was divided into three main categories (*content*, *coordination*, *other*) with one to four sub-categories each, and the transactivity dimension was divided into four main categories (*information sharing*, *quick consensus*, *transactivity*, *other*) with two sub-categories each (see the coding scheme presented in van Dijk et al., 2014, p. 361). A similar research context and adaptation of the Weinberger and Fischer (2006) coding scheme can be found in Gijlers et al. (2013).

The coding schemes developed as part of the research in the present thesis were to be applied to young learners collaborating in an embodied learning environment, where learners not only communicate through language, but with their bodies, their gestures, and their joint representation on the tablet screen. In such a co-present setting, communication is not represented in written text, but communication extends on bodily interactions with each other and the shared device, and only partially relies on language and oral communication. As in the case of van Dijk et al. (2014), the application of the coding scheme required appropriate adaptation to the specific research context and the sort of data that is emerging during the research (see also, Mercer, 2004).

Summarized, the specific research context (e.g., computer-mediated vs. co-present, adults vs. children) impacts to what extent an existing coding scheme can be utilized or requires adaptation to adequately sample dialogue qualities in the context of specific research questions (see also, Mercer, 2004). Table 6 and Table 7 provide a comparison of the original coding schemes by Weinberger and Fischer (2006) with the coding schemes developed in the course of this thesis. Throughout the comparison, it will become evident where changes to the original coding schemes were necessary to do justice to the specific research context of young co-present learners, and where the coding scheme could be adopted without changes. From the comparison in Table 6 and Table 7, it becomes also evident that the required changes were larger for the epistemic compared to the transactivity dimension. This may be explained by the assumptions that specific learning domains and contexts affect the epistemic dimension more than the (social) transactivity dimension which should be more context-independent. Pontecorvo and Girardet (1993) explain that because of "particular epistemological operations (e.g., types of explanation, ways of reasoning, conceptual frameworks), it is essential to take into account the peculiar features of each knowledge domain" (p. 367).

Table 6

Comparison of the Epistemic Dimension Coding Scheme by Weinberger & Fischer (2006) with the Epistemic Quality Coding Scheme Developed in the Course of the Present Thesis (published in Schmitt & Weinberger, 2019)

Epistemic dimension coding scheme by Weinberger & Fischer (2006)	Epistemic quality coding scheme developed in the course of the present thesis	Similarities and adaptions
	On-topic: regulation of the interaction	This category was added to refer to interactions regulating access to the technology or whose turn it is in speaking. These aspects are very frequent in co-present scenarios, but may be less frequent in online computer-mediated scenarios.
Construction of problem space	On-topic: concrete task-related regulation	These categories are very similar in both coding schemes with a focus on learners' talk about the concrete problem.
Construction of conceptual space	On-topic: abstract content-related regulation	Also here, the categories are very similar in both coding schemes, with a focus on learners' use of underlying theories.
Construction of adequate relations between conceptual and problem space; Construction of inadequate relations between conceptual and problem space; Construction of relations between prior knowledge and problem space	On-topic: procedural knowledge / strategies (also non-verbal)	Here, the coding scheme was adapted by collapsing three categories into one, which is largely due to the different target group (children instead of adults), and the co-present setting. Learners' verbalizations in the learning setting of this thesis tended to be very short and intimately tied to the bodily activities. These factors did not allow for a very differentiated analysis of the correct or incorrect connections between problem space and conceptual space and possible relations to prior knowledge. The presented category reflects and combines the ideas of the three original categories, in the way that prior knowledge is activated and underlying theories and concrete problems are connected whenever learners make use of strategies and demonstrate procedural knowledge, in other words, when they *apply* their knowledge.
Non-epistemic activities	Off-topic utterance	This category was adopted without changes. In both cases, it is referred to any off-topic talk.

Note. The aspects of the epistemic dimension (Weinberger & Fischer, 2006), and the epistemic quality (Schmitt & Weinberger, 2019) appear as they are named in the respective publications.

Table 7
Comparison of the Social Modes of Co-Construction Coding Scheme by Weinberger and Fischer (2006) with the Transactivity Coding Scheme Developed in the Course of the Present Thesis (published in Schmitt & Weinberger, 2019)

Social modes of co-construction coding scheme by Weinberger & Fischer (2006)	Transactivity coding scheme developed in the course of the present thesis	Similarities and adaptions
Externalization	Externalization	The externalization category was adopted without changes. In both coding schemes, externalization refers to a new utterance that does not build on something the learning partner previously said.
	Externalization as reaction	In addition to a simple externalization, the category externalization as reaction was added to capture instances where a learning partner reacted to the partner's contribution without it being any of the other categories, for instance, answering a simple question (for an example, see Table 9).
Elicitation	Elicitation	The categories in both coding schemes are very similar, with elicitation capturing the act of asking a question to the learning partner or provoking a reaction. In the coding scheme of this thesis, the intended reactions include also bodily activities. One change, however, was that elicitation was classified as a relatively low quality move (lower than quick consensus building) in the coding scheme by Weinberger and Fischer (2006), whereas in the coding scheme of the research in this thesis, it is regarded as a higher quality (higher than acceptance / refusal). Quick consensus building in computer-mediated communication and the coding scheme by Weinberger and Fischer (2006) may involve more than a simple "yes" or "no" for achieving that quick consensus, whereas in the co-present learning context of the research done in this thesis, a simple 'yes' or 'no' was very common. It was assumed that simply saying 'yes' or 'no' involves less effort and less consideration of the partner's contributions and abilities than asking a question. Thus, it was deemed justifiable to position elicitation as a category of higher quality in the adapted coding scheme.

Social modes of co-construction coding scheme by Weinberger & Fischer (2006)	Transactivity coding scheme developed in the course of the present thesis	Similarities and adaptions
Quick consensus building	Acceptance Refusal	Quick consensus building was adopted, but further differentiated in a simple acceptance vs. a simple refusal of the learning partner's contribution.
Integration-oriented consensus building	Integration	The integration category was adopted without changes. In both coding schemes, integration refers to attempts to rephrase and adopt the learning partner's contribution, going beyond a simple acceptance.
Conflict-oriented consensus building	Conflict-oriented consensus building	Also the highest transactive category, conflict-oriented consensus building, was adopted without changes, referring to attempts to challenge, change, or disagree with the learning partner's contribution, going beyond a simple refusal.

Note. The aspects of social modes of co-construction (Weinberger & Fischer, 2006), and transactivity (Schmitt & Weinberger, 2019) appear as they are named in the respective publications.

Table 8 and Table 9 display the final coding schemes of epistemic quality and transactivity with category descriptions and examples as they were used in studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4). To do justice to the context in which the verbalizations took place, the coders watched the videos while coding the transcriptions. In case that multiple categories would be applicable to one utterance, the highest one was selected.

Table 8

Coding Scheme Developed for Measuring Epistemic Quality, as Published in Schmitt and Weinberger (2019) / Study III

Category	Description of category	Examples
Off-topic utterance (OT)	Utterances not dealing with the iPad or the app	“I am hungry”
On-topic: regulation of the interaction (RegIA)	Dialogue, technology, task or learning partner are regulated; structuring of conversation without deeper meaning	“well…”, “Give me the iPad”, “Correct”, reading aloud a prompt
On-topic: concrete task-related regulation (TaskIA)	Hints or explanations that are very close to what one can see and perceive	“I need to stop here”, “this is 3 and this is 1”
On-topic: abstract content-related regulation (ContIA)	Hints or explanations that refer to abstract knowledge, more than what one can see or perceive	“double it”, “this is a third”
On-topic: procedural knowledge / strategies (also non-verbal) (Proc)	Strategies like counting, measuring, applying the prompts of the pedagogical agent	“1 2 3 4…”, “First estimate the height, then set the bars!”

Table 9
Coding Scheme Developed for Measuring Transactivity, as Published in Schmitt and Weinberger (2019) / Study III

Category	Description of category	Examples
Externalization (Ext)	Knowledge / thoughts are being externalized without reference to the learning partner	"There are no numbers anymore, it is always the same."
Externalization as reaction (ExtR)	Reaction to an elicitation without it being an acceptance, refusal, integration or conflict-oriented consensus building	A; "Where?" B: "Here, at the line"
Acceptance (Acc)	Simple and short acceptance of partners' statement	"Yes", "OK", "Good"
Refusal (Ref)	Simple and short rejection of partners' statement or action	"No", "Stop", "Wait"
Elicitation (Eli)	Asking a question or provoking a (verbal or activity-oriented) reaction	"The 3 needs to be higher, doesn't it?"
Integration (Int)	Elaborated acceptance incl. repetition or rephrasing of what's been said by the partner (no new content)	A: "Always half of it" B: "Always half, yes"
Conflict-oriented consensus building (COC)	Reference and modification, extension, relativization or counter argumentation to what has been said by the partner	A: "This needs to go higher" B: "You have to go lower. It needs to be three times the size"

According to Roschelle and Teasley (1995), language and action are closely intertwined, and actions may, for example, be used for externalizing concepts, or accepting the partner's suggestions by executing it physically. Likewise, in the presented coding schemes, verbal utterances can also target a bodily reaction, for instance, *elicitation* (see Table 9) not only targets a verbal answer, but may also suggest a concrete action within the app.

The categories of the transactivity coding scheme (see Table 9) can be mapped to Mercer's (1996) three categories of talk (disputational, cumulative, explorative; see section 4.1): *Externalization*, *externalization as reaction*, and *refusal* correspond to disputational talk, where learners merely try to get their own ideas across without engaging with the learning partner's ideas. *Acceptance* and *integration* correspond to cumulative talk, where learners add their knowledge together rather uncritically. *Conflict-oriented consensus building* and *elicitation* correspond to explorative talk, where learners actually engage in joint reasoning processes through critically challenging each other's contribution and, at best, converge on a shared and jointly created understanding.

While the different categories of epistemic quality and transactivity represent an ordinal scale (the categories are ranked from rather low to rather high quality), also the categories of comparatively low quality are important for collaborative talk. Even off-topic utterances (see the epistemic quality dimension, Table 8) serve an important social function in children's dialogue (Thornborrow, 2003). Externalizing an initial idea or a piece of knowledge, or asking the learning partner a question (see the transactivity dimension, Table 9) are the crucial first steps, making critical engagement with the content by the learning partner possible. An initial externalization also serves the function of restructuring one's own knowledge about the idea one wants to communicate, and can start a discussion (Weinberger & Fischer, 2006). Similarly, the *concrete task-related regulation* category (see the epistemic quality dimension, Table 8), may be of comparatively low epistemic quality, but serves the function of understanding and solving the tasks (see also, Weinberger & Fischer, 2006), which again can be considered as a first step in gaining a deeper understanding of the underlying principles. As such, all categories of epistemic quality and transactivity are important for productive talk, however, at best, learners would manage to go beyond pure externalizations and collaboratively reason together (see also Mercer, 1994, as cited in Falloon & Khoo. 2014). For concrete examples of the coding of learners' dialogue on these dimensions, see the excerpt displayed in Table 10.

Table 10

Example of Learners' Dialogue and the Coding of Epistemic Quality and Transactivity from the Transcribed Proportion Video Data (Translated from German to English)

Participant	Utterance	Epistemic quality	Transactivity
A	Um, I understand [unintelligible]	RegIA	Ext
B	This is three times as much, wait, this needs to go down. No, I think one more down.	ContIA	Eli
A	Almost. No, 1.5 it is after all.	TaskIA	COC
B	2.5	TaskIA	COC
A	Huh? Which one did you have? 1.5 or [unintelligible]? Wait! Like this	TaskIA	Eli
B	Wait, I pull this down a bit. No, that's only 1. Wait, I will also pull down this one a bit, then it will be easier	TaskIA	Ref
A	No, this [unintelligible] not.	RegIA	Ref
B	On the second line	TaskIA	Eli

For the data analysis of the research in this thesis, specifically for studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4), the coding schemes for epistemic quality and transactivity had to be trained to be operable by multiple coders (three student assistants and I), in order to reach sufficient inter-rater reliability values (Krippendorff's alpha; Krippendorff, 2012). The alpha values for epistemic quality and transactivity, assessed for 7 % of the analyzed data, were good with $\alpha = .88$ for epistemic quality and $\alpha = .77$ for transactivity (Schmitt & Weinberger, 2019 / study III, section 7.3).

6.2.4. Conclusion

To better understand the ongoing learning processes of collaborative learning with a jointly used device, one needs to look at the processes from multiple angles: For example, to what extent are learners continuously engaged in the activity? What quality is the content of learners' talk? To what extent do the learners refer to each other when they speak? Which sort of emotions do they express and how frequently? Investigating the processes during collaborative learning, and to what extent they depend on different learning environment designs, cannot be determined through one variable alone. By taking into account multiple relevant process variables, the presented coding schemes aim at providing a more exhaustive and differentiated view of the processes involved during collaborative embodied learning. The past sections gave an overview on prior research and analysis methods for the selected variables. While this prior work is a highly valuable source for conceptualizing and operationalizing these variables, it was necessary to move beyond the existing work and develop own coding schemes. On one hand, some of the related methods were very qualitative in nature (e.g., Mercer, 1996), on the other hand, the methods were usually created for a specific learning context (e.g., adults' computer-mediated CSCL; Noroozi et al., 2013). These two aspects imply that these methods cannot be readily adopted, but there was a need to develop standardized coding schemes which are tailored to the specific learning context of children collaborating in an embodied learning environment. Each of the process variables is relevant to learning (see chapter 4) and a qualitative way of analysis is required to capture them; for instance, it would be barely possible to assess the epistemic quality in dialogue by means of a questionnaire.

The presented process variables can be separated as demonstrated with the different coding schemes, however, there are related to each other. For instance, positive emotions may help to sustain on-task behavior, respectively, engagement may have an affective component (Deater-Deckard et al., 2013). In a model proposed by Polo, Lund, Plantin, and Niccolai (2016), momentary emotions, the overall emotional tone of a discussion, and the type of talk that learners are engaged in are all related and, for example, an acute emotional experience may impact the overall tone, and the overall tone of the discussion elicits different discussion styles, like cumulative talk (in accordance with the talk types by Mercer, 1996). Hence, it is important to look at the different facets of a collaborative situation simultaneously to get a fuller picture of what is going on.

As suggested in the *eSOCS* coding scheme (Deater-Deckard et al., 2011), coding of process variables can be done for the individual learner also in a collaborative setting. This approach was also taken in the analysis in this thesis, specifically studies III (Schmitt & Weinberger, 2019; section 7.3) and IV (Schmitt et al., in preparation; section 7.4): First, the variables were coded for each individual of a dyad, and in the next step, a group value was calculated by aggregating the individual values to a group mean.

6.3. Overall Conclusion: Research With Children

In the present thesis, methods for research with children are developed. On one hand, working towards more child-friendly survey instruments for young learners; on the other hand, working towards creating and refining coding schemes to qualitatively capture learning processes on multiple dimensions, such as quality of dialogue, emotions, or focus on the task.

This chapter on research with children will be concluded with a comment on the question of qualitative vs. quantitative research. Both qualitative and quantitative approaches are useful and indeed are being utilized in social sciences studies (Mercer, 2004). In addition to quantitative methods, such as a closed item questionnaires, in the course of this research, coding schemes for measuring verbal and non-verbal qualities of the learning processes were developed and utilized as a *qualitative* approach towards capturing collaborative learning facets as they unfold during the learning experience. This approach of coding and counting behavior or talk according to predefined categories may be classified as *quantitative* research by some researchers, implying simplification, data reduction, and a loss of flexibility (Mercer, 2004). However, one may also refer to this approach as “quantifying qualitative analyses” (Chi, 1997, p. 271; see also Vogel & Weinberger, 2018). The original data was fairly rich, i.e., transcribed video recordings, allowing for multimodal analysis of processes, for example, coding dialogue on two dimensions with in total 12 categories. Only in a second step, that data was aggregated and transformed into numbers for statistical analyses. In addition to this quantified data, excerpts of learners’ talk are presented (Schmitt & Weinberger, 2019 / study III; Schmitt et al., in preparation / study IV). Providing and commenting on these excerpts goes beyond coding and counting and enables the reader to get a better idea of the qualities of the learning processes. Or as Mercer (2004) puts it: “The examples of talk provided to any audience for the research are real: they are not asked to take on trust the validity of an abstracted categorization scheme.” (p. 142). In essence, both are important: Coding and counting large amounts of data, and concretely illustrating learners’ processes by presenting and discussing exemplary cases. The first allows for statistical comparisons between experimental groups and testing hypotheses, the latter for illustrating the processes in depth (see also, Schmitt & Weinberger, 2019 / study III; van Boxtel et al., 2000).

This chapter concludes the theoretical and methodological introduction to the research done in the present thesis. The following chapter 7 will give a summarized overview on the four studies of this thesis.

7. The Four Studies of the Present Thesis

The past six chapters discussed studies, theories, and methods regarding this thesis' core topic of children's co-present collaborative and embodied learning with a shared tablet. For researching this core topic, one needs to have a good understanding of co-present CSCL, and of theoretical perspectives like embodied cognition to understand bodily learning with a co-present learning partner. Further, it is essential to be aware of critical learning processes on multiple dimensions and ways of operationalizing them. These matters were covered along with an introduction of the domain proportional reasoning and the *Proportion* learning environment, in order to motivate the four studies of this thesis and to integrate them into a larger research context. Next, chapter 7 will give a summarized overview on these four studies. The studies are internationally published, or in preparation. Study I (Schmitt & Weinberger, 2018; section 7.1) provides an overview on learning scenarios and research regarding CSCL in primary and secondary education; study II (Schmitt et al., 2019; section 7.2) presents work on developing an instrument for better surveying young learners; study III (Schmitt & Weinberger, 2019; section 7.3) demonstrates results on the influence of two sorts of prompts on learning processes and outcomes in the *Proportion* learning environment; study IV (Schmitt et al., in preparation; section 7.4) zooms in on learning process heterogeneity to better understand and untangle the conditions of successful collaborative embodied learning.

7.1. Study I: Computer-Supported Collaborative Learning: Mediated and Co-Present Forms of Learning Together[3]

Study I is published as a book chapter in the international and peer-reviewed handbook *Second Handbook of Information Technology in Primary and Secondary Education*, and gives an overview on co-present and computer-mediated learning scenarios and research findings regarding CSCL in primary and secondary education, along with associated challenges and ways to support learning in these scenarios.

---

[3] Study I is published as:
Schmitt, L. J., & Weinberger, A. (2018). Computer-supported collaborative learning: Mediated and co-present forms of learning together. In J. Voogt, G. Knezek, R. Christensen, & K.-W. Lai (Eds.), *Second handbook of information technology in primary and secondary education* (pp. 217–231). Cham, Switzerland: Springer. doi:10.1007/978-3-319-53803-7_15-1

7.1.1. CSCL

CSCL involves peers learning together with different sorts of computer-support. Collaborative learning can be distinguished from cooperative learning, with the former focusing on learners working closely together at a shared task (Roschelle & Teasley, 1995), and the latter rather implying that learners may split a larger task into parts which can be completed by individual learners. Collaborative learning can be characterized by high degrees of transactivity (Noroozi et al., 2013; Roschelle & Teasley, 1995), interactivity (Dillenbourg et al., 2009), and knowledge convergence (Weinberger, Stegmann, & Fischer, 2007), and that the participants are striving towards common ground (Beers, Boshuizen, Kirschner, & Gijselaers, 2006). For children in particular, also mutuality and intimacy are important in collaborative learning (Crook, 1998). Learners' processes towards knowledge acquisition during a collaborative activity can be observed on several dimensions of quality, for instance, regarding argumentative or epistemic quality, or the degree of transactivity (Weinberger & Fischer, 2006). One of the main research interests in CSCL is how to support learners in engaging in these productive learning processes, for example, through scaffolding or scripting (e.g., Noroozi et al., 2013).

7.1.2. Computer-mediated vs. co-present computer-supported collaborative learning: Scenarios and challenges

Computer-mediated CSCL scenarios cover scenarios where learners are at different places and communicate through online platforms. Communication may be synchronous or asynchronous, and may be more or less rich, for example, text-only vs. video communication. For instance, learners can build and exchange arguments (e.g., text-based in a discussion board), collaborate on discovering scientific phenomena in massive open online courses (MOOCs), share and co-construct knowledge in social media (e.g., Facebook), or interact in richer environments, such as shared whiteboards, video conferences, or virtual worlds (e.g., Ertl, Fischer, & Mandl, 2006; Ligorio & Van der Meijden, 2008; Tsovaltzi et al., 2015). Known challenges of computer-mediated CSCL scenarios span the lack of social context cues, respectively a higher need for coordination, negative group phenomena like free-riding, lack of motivation, or high drop-out rates (e.g., Jordan, 2015; Van der Meijden & Veenman, 2005).

Co-present CSCL scenarios cover scenarios where the learners are physically together and can directly communicate. Technologies for these scenarios include, for instance, whiteboards, tablets, or handhelds. More recent technologies, for instance multi-touch screens, allow for equal access to the device by multiple learners, and may support embodied learning (e.g., Schneps et al., 2014). Examples for co-present CSCL scenarios in primary education are environments where learners embody particles (Danish et al. 2015), discuss together at a tabletop (Mercier et al., 2014), or solve math problems with handheld devices (Roschelle et al., 2010); in secondary education, learners may, for example, inquire topics such as global warming on the web (Raes, Schellens, De Wever, & Benoit, 2016). Known challenges of these co-present scenarios include questions of territoriality and dominance (e.g., Moed et al., 2009; Rick et al., 2015), or overly playful behavior in rich and engaging learning environments (Danish et al., 2015).

7.1.3. Facilitating learning in CSCL scenarios

There are two basic ways to facilitate learning in CSCL scenarios: Structuring and regulating (Dillenbourg, 2002).

Structuring may be realized by means of CSCL scripts that specify roles and activities (e.g., Scheuer, Loll, Pinkwart, & McLaren, 2010), usually before the start of the learning activity. One example for a successful script is the *Guided Reciprocal Peer Questioning* script (Gelmini-Hornsby, Ainsworth, & O'Malley, 2011), but scripting approaches also bear the danger of *over-scripting* learners, potentially spoiling productive natural interaction patterns (Dillenbourg, 2002). Scaffolding is a related approach to support learners in solving the tasks, and to increase their understanding, so that future tasks can be solved independently (Reiser, 2004). As with scripts, scaffolding approaches can have positive effects, for example on performance, but also costs, such as reduced motivation (Chen & Law, 2016; see also, Schmitt & Weinberger, 2019 / study III, section 7.3).

Regulating refers to regulation of the ongoing interactions, for example, with awareness tools that provide information on the group processes, such as levels of participation. Also awareness tools showed positive effects on learning-relevant variables, such as transactivity or cohesion (e.g., Gijlers et al., 2013; Phielix, Prins, & Kirschner, 2010), but overall, their usefulness is dependent on learners' ability to productively use the awareness information (Janssen et al., 2011).

7.1.4. Conclusion

Learning in CSCL scenarios comes in many forms and, in spite of the strengths of CSCL, produces specific challenges which may be addressed by structuring and regulating learners. Further research may focus on how to improve the effectiveness of awareness tools, for instance by means of scripting (e.g., Janssen et al., 2011), and on questions of orchestrating and bridging different learning scenarios and providing appropriate support throughout the different learning arrangements.

7.2. Study II: BiCo: A Bipolar Continuous Rating Scale for Children's Technology Evaluation[4]

Study II is published in the international and peer-reviewed journal *Technology, Pedagogy and Education*. The focus lies on a review of existing instruments to survey children's evaluation of technologies, and to describe and justify the design of a new instrument (BiCo) to address the identified problems. The newly developed BiCo instrument is then piloted and compared with the established Smileyometer (e.g., Read, 2008) instrument.

7.2.1. Theoretical background: Empirical research with children, questionnaires for children, and developing the BiCo rating scale

Doing research with, and not just about, children (e.g., Fuchs, 2005; Scott, 2000) is a crucial step for the development of good educational technologies that enable learning and are at the same time enjoyable to use for young learners. Beyond participatory design of technologies or formative evaluation methods with usually a small number of participants (Nielsen, 1993), there is also a need for summative evaluation methods with a larger sample for enabling hypothesis testing and drawing quantitative conclusions (Hall, Hume, & Tazzyman, 2016). Thus, children's self-report of their experience comes into focus (Sim et al., 2006). While traditionally children's self-reporting abilities have attracted doubts (Fuchs, 2005), it is nowadays acknowledged that surveying children via self-report (questionnaires) is valuable and reasonable, but there is a need for designing appropriate surveying methods particularly for children (e.g., Bell, 2007).

Challenges for children's questionnaire use are ambiguous and complex questions (de Leeuw & Otter, 1995), negative wording (inverted items) (Marsh, 1986), or a peer groups' influence when they are nearby (Scott, 2000). Thus, the challenges are similar to the ones of adults, but possibly intensified as the children's cognitive and social abilities are not yet fully developed (Borgers, de Leeuw, & Hox, 2000).

Rating scales are one type of closed items to easily measure and quantify participants' subjective evaluations. Rating scale design needs to take care of several considerations (Moosbrugger & Kelava, 2008): Continuous vs. discrete answering options, uni- vs. bipolar format, symmetric vs. asymmetric scale, and denomination (numerical, verbal, symbolic). The

[4] Study II is published as:
Schmitt, L. J., Rick, J., & Weinberger, A. (2019). BiCo: A bipolar continuous rating scale for children's technology evaluation. *Technology, Pedagogy and Education, 28*(5), 503–516. doi:10.1080/1475939X.2019.1661279.

Smileyometer instrument (Read, 2008, see Figure 7) is one such (discrete, uni-polar, asymmetric, symbolic) rating scale developed particularly for children, utilizing five smiley faces. Smileyometer is widely used in child-computer interaction research, but a reoccurring issue is the opinion ceiling effect (e.g., Looije, Neerincx, & de Lange, 2008; Johnson et al., 2016), when all responses are highly positive across items and experimental conditions, not allowing for any differentiation (solely high item difficulty; Moosbrugger & Kelava, 2008).

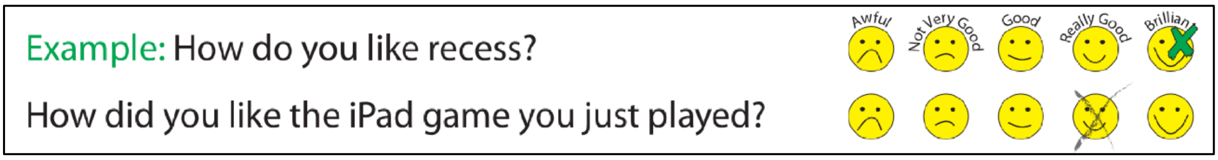


Figure 7. The Smileyometer instrument (e.g., Read, 2008) as used and published in Schmitt, Rick, and Weinberger (2019) / study II. One example item, and one survey item (translated from German to English).

The opinion ceiling effect can be explained by several contributing factors: First, educational technology subjected to empirical studies usually is novel and of good quality, so children usually actually like it. Second, younger children tend to give more extreme (Chambers & Johnston, 2002) and more extremely positive ratings (Read & MacFarlane, 2006). Third, acquiescence bias (Colombo & Landoni, 2014) or compliance towards the researchers (Bell, 2007) may lead to overly positive responses. Fourth, satisficing strategies, specifically when questionnaires are long (Borgers et al., 2000), may lead to a response behavior where choosing the most positive answer for all items seems to be the easiest way. Smileyometer already addresses many of the challenges of surveying children, for example, the symbolic representation is supposed to be more accessible and enjoyable for children, and the asymmetry of the scale towards more happy faces might attenuate highly positive responses (a Smileyometer adaptation called *Five Degrees of Happiness* (Hall et al., 2016) with only happy faces resulted in higher response variance). However, other common techniques to combat, for example, the acquiescence bias, like presenting inverted items (Moosbrugger & Kelava, 2008), cannot be addressed with the Smileyometer, and are problematic for using with children anyway (Bell, 2007).

An alternative approach for surveying children may be to utilize their strengths in drawing comparisons. Instruments that require to compare several technologies or variants of a technology simultaneously, such as *Fun Sorter* and *Again Again* (Sim & Horton, 2012), or *This or That* (Zaman et al., 2013), seem to work better with young respondents. For example, children aged 7-8 years old could give a differentiated response for different evaluation criteria (fun, ease of use, learning) when they compared three different software applications (Sim et al., 2006). However, the comparison approach entails that each participant engages with each technology in a within-subject design, in order to draw a comparison between them, which is not always sensible, depending on the research question. See Table 11 for a comparison of different survey instruments for children.

BiCo (Bipolar Continuous Rating Scale) was developed to address the present challenges of surveying children by harnessing their ability for drawing comparisons, and giving them the opportunity for a fine-grained judgement by means of the continuous nature of the scale. As can be seen in Figure 8, BiCo presents two opposing statements, and the respondents can indicate on a continuous axis to what extent they agree to one or the other side. Summarized, the following design choices informed BiCo's development:

- The **bipolarity** of the instrument enables the children to draw relative comparisons instead of merely agreeing to each item. Moreover, the bipolarity allows for switching the expected ('positive') answer within a series of items to prevent satisficing strategies. Smileyometer always has the most positive answer option on the right side. Randomly switching the side of the expected answer is assumed to achieve a similar effect as inverted items, but without the cognitive burden of processing negatively formulated items.
- The **continuous** axis may afford more fine-grained responses, as a respondent can indicate high agreement with a statement without resorting to the most extreme answering option. Further, the continuity does justice to most statistical tests assuming metric scales, which is only approximated by Likert like scales.
- The **verbal** denomination is assumed to evoke more consistent interpretations of the items across respondents.

Table 11
Overview on and Comparison of Available Survey Instruments for Children, as Published in Schmitt, Rick, and Weinberger (2019) / Study II

Instrument	How it works	Advantages	Disadvantages
Smileyometer	Symbolic scale with 5 smiley faces from 'awful' to 'brilliant'	- Enjoyable and easy to use - Asymmetric scale	- Often associated with the opinion ceiling effect - Cannot include inverted items - Not well suitable for concepts beyond 'fun'
Fun Sorter	Ranking of the multiple technologies/applications that the child has interacted with (different dimensions, e.g. fun)	- Drawing comparisons is accessible for children - Avoiding the opinion ceiling effect	- Only suitable for the simultaneous evaluation of multiple technologies/applications (within-subjects design)
Again Again	Table in which the child indicates if it would want to use the technology/application again ('yes', 'maybe', 'no')		
This or That	Child indicates which technology/application they favor (different questions, e.g. which one would they want to use again)		

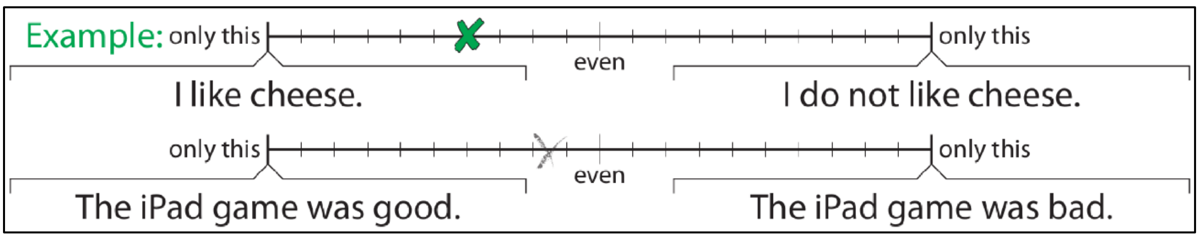


Figure 8. The newly developed BiCo instrument as used and published in Schmitt, Rick, and Weinberger (2019) / study II. One example item, and one actual item (translated from German to English).

Summarized, by letting the children draw comparisons on a bipolar and continuous axis, Bico is hypothesized to mitigate the opinion ceiling effect compared to Smileyometer. Inverting BiCo's scale may mitigate satisficing strategies, and the verbal labels are assumed to evoke more consistent interpretations of the items across participants. Further, BiCo's applicability is broader compared to Smileyometer, which focuses on the concept of fun but has limited applicability regarding other concepts. Thus, BiCo allows for covering a wide range of concepts while sticking to the same instrument, making the answering process easier for the young respondents. A last advantage of BiCo is its usability in between-subjects studies where every participant only gets to interact with one technology.

BiCo items were created and used (for instance, see Figure 8), along with their Smileyometer counterparts (for instance, see Figure 7), in a study about fourth graders mathematics learning with tablets (see the following section 7.2.2). When creating the items, the focus was on conceptual equivalence, as well as positive and simple wording. The item construction process underwent several iterations of revisions within the research team as well as with an external expert.

7.2.2. A study comparing BiCo and Smileyometer: Background, methods and results

In one study with the *Proportion* app (in total n = 125 participants, mean age: 10 years and 3 months, *SD* = 6.5 months), three experimental conditions (collaborative learning in multi-touch mode, collaborative learning in single-touch mode, individual learning) were investigated. The participants individually filled in a pre-questionnaire and pre- math test, worked with the *Proportion* app for 40 minutes, and filled in a post-questionnaire and post- math test. The pre-questionnaire focused on demographical aspects as well as attitudes towards school, math, and collaborative learning (assessed with Smileyometer); the post-questionnaire focused on aspects of the app, aspects of the collaborative experience, and subjective learning gain (assessed with Smileyometer). After collecting the first data (n = 43), the opinion ceiling effect with Smileyometer was evident (e.g., means between 4.29 and 4.67 (out of 5) in the post-questionnaire), motivating the development of BiCo. BiCo was then employed alongside Smileyometer for the further data collection and analysis, with the goal of providing first data comparing the two instruments, and to inform us to what extent BiCo may be a useful instrument to capture children's technology evaluations. (The original three experimental conditions are of no further relevance for the BiCo vs. Smileyometer analysis, and therefore, will not be elaborated on.)

In the subsequent data collection (n = 82, mean age: 10 years and 5 months, *SD* = 6.6 months), the post-questionnaire items were presented in both formats (BiCo and Smileyometer) to enable a within-subject statistical testing. The expected answers of the BiCo items were randomly switching sides (inverting the scale) with the goal of reducing acquiescence bias. For an overview on the items of the study, see Table 12.

Table 12
Items Used in the Study in both Smileyometer and BiCo Format (Translated from German to English), as Published in Schmitt, Rick, and Weinberger (2019) / study II

#	Smileyometer Questions	BiCo Statements to Compare	
(a)	Would you like to continue?	I would like to stop.	I would like to continue.
(b)	Did you learn something about math with the iPad?	I learned a lot using the iPad.	I did not learn anything using the iPad.
(c)	Did you help your partner a lot?	I helped my partner a lot.	I did not help my partner.
(d)	How helpful was your partner?	My partner distracted me.	My partner was helpful.
(e)	How well did you work together?	We worked well together.	We would have been better off working alone.
(f)	How did you like the iPad game you just played?	The iPad game was bad.	The iPad game was good.

For the analysis, the BiCo and Smileyometer responses were converted into a value between 0 and 1, with 0 meaning no agreement with a statement, and 1 meaning full agreement with a statement.

To test if the items in both formats measured the same concepts, correlations were calculated. All correlations were positive and significant, with some medium-sized correlations (item (b): r (82) = .482, p = .000; item (c): r (64) = .479, p = .000; item (d): r (65) = .325, p = .008; item (e): r (65) = .339, p = .006), and some high correlations (item (a): r (80) = .681, p = .000; item (f): r (82) = .601, p = .000).

To test a possible mitigation of the opinion ceiling effect, repeated measures ANOVAs with the item format as within-factor were conducted. Using a Bonferroni correction, the alpha-level was set to .008. As can be seen from the descriptive data in Table 13, the means were lower for all items in Bico vs. Smileyometer format. Though absolute differences were small, the difference was statistically significant for three out of six items: Regarding item (a), there was a large effect with F (1, 79) = 31.718, p = .000, η^2 = .286. Regarding item (b) (F (1, 81) = 10.251, p = .002, η^2 = .112) and item (e) (F (1, 64) = 8.342, p = .005, η^2 = .115), there were significant differences with medium-to-large effect sizes. There were no significant differences depending on item format for items (c), (d), and (f).

Table 13
Means and Standard Deviation for the Items in Smileyometer and BiCo Format, as Published in Schmitt, Rick, and Weinberger (2019) / Study II

#	Item	n	Smileyometer		BiCo	
			M	SD	M	SD
(a)	I want to continue	80	.844	.250	.709	.283
(b)	I learned something	82	.720	.253	.642	.275
(c)	I helped my peer	64	.688	.264	.638	.241
(d)	My peer helped me	65	.785	.225	.782	.195
(e)	We collaborated well	65	.885	.183	.803	.211
(f)	I liked the game	82	.762	.257	.732	.255

Moreover, the frequency of selecting the most extreme answer option for an item was lower for BiCo (3.1% – 13.6%, average of all items: 9.1%), compared to Smileyometer (29.2% – 67.7%, average of all items: 46.9%), see Table 14.

Table 14
Percentage of Extreme Answers with Smileyometer vs. BiCo, as Published in Schmitt, Rick, and Weinberger (2019) / Study II

#	Item	Smileyometer		BiCo	
		n	Percentage	n	Percentage
(a)	I want to continue	80	67.5%	82	8.5%
(b)	I learned something	82	30.5%	82	4.9%
(c)	I helped my peer	65	29.2%	65	3.1%
(d)	My peer helped me	66	42.4%	65	12.3%
(e)	We collaborated well	65	67.7%	66	13.6%
(f)	I liked the game	82	43.9%	82	12.2%

In sum, the statistical analysis indicates that 1) both item formats correlate positively, supporting the assumption that they still measure the same concept, 2) BiCo produces comparatively lower means, statistically significant for three out of six items, reducing the opinion ceiling effect, and 3) the behavior of selecting the most extreme answer was less frequent for BiCo than for Smileyometer items.

7.2.3. Discussion

This research introduced the challenge of adequately surveying children and presented common questionnaire instruments along with their strengths and weaknesses, motivating the need for developing a novel instrument to better capture facets of children's technology evaluation. BiCo was designed to overcome the opinion ceiling effect by utilizing children's ability to draw relative comparisons and presenting two opposing statements which can be compared and agreed to on a continuous scale. Moreover, BiCo's design allows for its usage in between-subjects experimental settings where each participant is only evaluating one (and not multiple) technologies. Also, the scale can be inverted in order to reduce the acquiescence bias, which is a similar approach to negatively formulated items, but without the additional cognitive burden of negative wording. Finally, the design allows for a broad applicability of the instrument, going beyond the concept of fun.

In the statistical analysis, all BiCo items correlated positively with the Smileyometer items, and produced lower means, descriptively for all items and statistically significant for three out of six items, supporting the assumption that BiCo mitigates the opinion ceiling effect while still measuring the same concepts. Also, learners' behavior of choosing the most extreme answer was less frequent with BiCo. These findings support BiCo's utility as a new instrument in child-computer interaction research.

An identified weakness of this research is that the children used both instruments in the post-questionnaire, which may have influenced their response behavior. However, in favor of the chosen approach is that item order effects are less problematic with young questionnaire respondents (Fuchs, 2005), and that within-subject designs reduce error variance.

Study II ends with an outlook on future research: For properly establishing BiCo as an instrument, more research needs to be done, for example, investigating BiCo's utility with children from different backgrounds (different ages; different educational, social, and cultural backgrounds) and in different learning settings. Moreover, future research may look into the mental effort needed to respond to BiCo items compared to other instruments. Finally, one could investigate different designs of BiCo, for instance, a design with more exhaustive labels (Borgers, Hox, & Sikkel, 2003), and explore the topic of how to create appropriate bipolar statements.

7.3. Study III: Fourth Graders' Dyadic Learning on Multi-Touch Interfaces – Versatile Effects of Verbalization Prompts[5]

Study III is published in the international and peer-reviewed journal *Educational Technology Research and Development*. The focus is on investigating the effects of additional prompts on learning processes and outcomes in the *Proportion* learning environment.

7.3.1. Theoretical background: Multi-touch interfaces for collaborative embodied learning, scaffolding CSCL, and proportional reasoning

Multi-touch devices may be utilized for collaborative learning activities (e.g., Mercier et al., 2014), and support embodied learning (e.g., Schneps et al., 2014). To better understand collaborative learning with a multi-touch device, learners' multifaceted learning processes need to be taken into account, including cognitive and emotional learning processes, which become observable through verbal and bodily process data. Emotions, task focus, and the quality of learners' dialogue are all crucial variables determining successful learning (e.g., Baker & Lund, 1997; Pekrun et al., 2002; Weinberger & Fischer, 2006). However, procedural and conceptual challenges are frequent in CSCL (Furberg, 2016). In study III, the *Proportion* tablet app is enriched with two sorts of prompts: Strategy prompts (STRAT) to address procedural challenges, and verbalization prompts (VERB) to address conceptual challenges.

Scaffolding for strategic behavior may be achieved by breaking down a complex task, by suggesting a reasonable sequence of activities, or by requesting the learner to be strategic and make up a plan before starting to work (Jackson et al., 1998; Schukajlow, Kolter, & Blum, 2015; Sharples et al., 2015). This added support may improve learners' confidence of success (Belland, Kim, & Hannafin, 2013), and was found to foster expert-like problem solving (Schoenfeld, 1987), as well as mathematical modelling competency (Schukajlow et al., 2015).

While certain qualities of learning processes are highly relevant for effective learning, such as building and sharing knowledge (Baker & Lund, 1997), learners may not be focused on the task (Danish et al., 2015), or may not engage in productive dialogue (Falloon & Khoo, 2014). Scaffolding for verbalizations is intended to support learners to elaborate (King, 1990), to stay focused, and to enhance their learning (Reiser, 2004). This type of scaffolding may request the learner to explain and discuss their reasoning, for instance, with verbalization prompts like ''Explain why…'', ''How are…and…similar?'' (King, 1990, p. 669). Prior research showed

[5] Study III is published as:
Schmitt, L. J., & Weinberger, A. (2019). Fourth graders' dyadic learning on multi-touch interfaces – versatile effects of verbalization prompts. *Educational Technology Research and Development, 67*(3), 519-539. doi:10.1007/s11423-018-9619-5.

that prompting for verbalizations has positive effects on students' understanding of the learning material (King, 1990), and fosters an interactive and elaborate discussion (Gelmini-Hornsby et al., 2011), hence, positively impacts transactivity and epistemic quality.

Table 15 gives an overview about the two prompt types as implemented in the study and their theoretical foundations. The STRAT prompts were designed to evoke thinking and planning activities, the VERB prompts targeted collaborative reflection and abstraction. Figure 4 displays how the prompts were integrated into the app interface with one example each for VERB and STRAT prompts.

While scaffolding generally showed to have positive effects on learning, it may also have negative effects by interrupting learners' flow, resulting in decreased motivation and learning outcomes (Wouters et al., 2015). Little is known about the effects of scaffolding in learning environments like *Proportion* that are immersive and involve learners with bodily activities. So, while positive effects may be assumed based on previous literature, scaffolding may also disrupt learning in this particular context.

The learning domain of *Proportion* is proportional reasoning, a challenging mathematics domain (e.g., Boyer et al., 2008). Previous research suggests that learning environments which include bodily experiences with proportionality may be effective in supporting students in their proportional reasoning (e.g., Reinholz et al., 2010). Following these approaches, the *Proportion* learning environment utilized in the present study affords bodily activities and active manipulation of the tablet screen through the fingers (Rick, 2012a).

Table 15
Overview on Strategy and Verbalization Prompts, as Published in Schmitt and Weinberger (2019) / Study III

Prompt type	Level	Prompt text	Theoretical foundation
STRAT	1, 4, 7, 10, 13, 16, 19	A: Tip for all tasks: What is higher, orange or blue? First say it out loud, then you do!	Guiding learners by providing additional structure and fostering strategic behavior that can applied throughout the learning phase (Jackson et al., 1998; Schoenfeld, 1987; Schukajlow et al., 2015; Sharples et al., 2015)
	2, 5, 8, 11, 14, 17, 20	B: Tip for all tasks: First think and provide an estimate, then set the bars' correct height!	
	3, 6, 9, 12, 15, 18, 21	C: Tip for all tasks: If the task is hard and you're stuck, what might help is to discuss and talk!	
VERB	1, 4, 7, 10, 13, 16, 19	A: Explain to your learning partner: What did one need to do in order to solve the task?	Supporting knowledge construction by eliciting high-quality verbalizations, requesting concrete explanations to pre-determined questions (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001; King, 1990; Reiser, 2004)
	2, 5, 8, 11, 14, 17, 20	B: Describe to your learning partner: What could one learn in this task?	
	3, 6, 9, 12, 15, 18, 21	C: Explain to your learning partner: What do all of these tasks have in common?	

In study III, it was hypothesized that both prompts separately and their combination will contribute to a higher task focus (hypothesis 1). Moreover, STRAT prompts were assumed to further positive emotions and VERB prompts to increase negative emotions (hypothesis 2). VERB prompts were hypothesized to enhance the quality of dialogue (transactivity and epistemic quality) (hypothesis 3). Finally, both prompts separately and their combination were assumed to further learning gains (hypothesis 4).

7.3.2. Methods

In a 2 × 2 factorial design with pre- and post-test measures, data from $n = 162$ fourth graders were collected. The participants were on average 10 years and 4 months old ($SD = 6.7$ months); 50% were female. Pupils from different school backgrounds (more or less likely to proceed to Gymnasium in fifth grade; Regionalverband Saarbrücken, 2012) were included. The participants were randomly assigned to their learning partner and one of the experimental conditions: STRAT – only strategy prompts, VERB – only verbalization prompts, VERB-STRAT – both prompts, control – no prompts. After individually completing a first survey and a pre- math test, the participants collaboratively used *Proportion* for 40 minutes. This learning phase was video-taped. After the learning phase, there was a 5 minute break, followed by individually filling in another survey and a post- math test. The whole experiment cycle lasted approximately 90 minutes and replaced two regular school lessons.

Regarding the outcome variable *learning gain*, a math test on proportional thinking was adapted from Rick, Rogers, Haig, and Yuill (2009), and administered prior and after the learning phase. This test consisted of near and far transfer tasks. For the final analysis, we eliminated 7 out of 28 items based on an item analysis. Cronbach's alpha in the revised version was .76 for near transfer and .68 for far transfer.

Regarding the measured process variables, we selected specific samples of the video material: From each second level (levels 1, 3, 5, …21), we selected every second problem. This procedure was chosen in order to reduce the large amount of data, while sampling the interactions throughout the whole learning period. Coding of the non-verbal process variables *positive emotions* and *negative emotions* (Table 3), and *task focus* (Table 5) focused on eye contact for task focus (conceptualized as *off-task behavior*), and gestures for emotions (e.g., Deater-Deckard et al., 2014; Lehman et al., 2008). The dialogue process variables *epistemic quality* and *transactivity* were measured based on transcriptions of the video samples. Upon classifying the individual utterances into the different categories of *epistemic quality* and *transactivity* (Table 8 and Table 9), we built two global scores with the aggregated and weighted categories. Weighting was based on the assumption that the different categories represent an ordinal scale (Berkowitz & Gibbs, 1983; Teasley, 1997; Weinberger & Fischer, 2006), and that the weights would do justice to their presumed relative quality. Inter-rater reliability was given for all qualitative variables with Krippendorff's alpha (Krippendorff, 2012) values of $.77 \leq \alpha \leq .92$. Detailed coding schemes for all process variables can be found in chapter 6.2.

7.3.3. Results

Regarding the non-verbal process variables, we in total observed 805 instances of off-task behavior, 274 positive emotions, and 232 negative emotions throughout the entirety of the analyzed sample. Regarding off-task behavior, a two-factorial ANOVA revealed no statistically significant main effect of STRAT (F (1,76) = 0.159, p = 0.691), but a highly significant main effect of VERB (F (1,76) = 18.190, p = 0.000, η^2 = 0.19), indicating that off-task behavior was increased through verbalization prompts. The interaction effect of STRAT and VERB on off-task behavior was not significant: F (1,76) = 2.451, p = 0.122. Regarding emotions, we found no significant effect of STRAT on positive emotions (F (1,76) = 1.919, p = 0.170), but a statistically significant effect of VERB on negative emotions (F (1,76) = 7.019, p = 0.010, η^2 = 0.09), indicating that the frequency of negative emotions was increased through verbalization prompts. See Figure 9 for the effects of VERB and STRAT on the non-verbal variables.

Regarding the dialogue process variables, we in total observed 3564 dialogue turns, see Table 16 for a descriptive statistics regarding the coded categories of dialogue quality. Regarding both transactivity (F (1, 70) = 7.241, p = 0.009, η^2 = 0.094), and epistemic quality (F (1,70) = 9.437, p = 0.003, η^2 = 0.119), there were significant main effects of VERB, indicating that the VERB prompts improved both aspects of dialogue quality, see *Figure 10* and *Figure 11*. The full article (Schmitt & Weinberger, 2019) also presents an illustrative case analysis showing how the VERB prompt successfully stimulates a dyad to engage in a reflection phase after solving a problem, compared to an unprompted dyad which proceeded to the next task without taking their time to reflect.

Finally, regarding learning gain, there was a general improvement in near transfer tasks independent from conditions (F (1,77) = 11.179, p = 0.001, η^2 = 0.13), but no significant interaction with either of the prompts (VERB × point in time: F (1,77) = 0.061, p = 0.805; STRAT × point in time: F (1,77) = 0.585, p = 0.447), or their combination (VERB × STRAT × point in time: F (1,77) = 1.502, p = 0.224). Concerning far transfer tasks, there was neither a general improvement (F (1,77) = 0.174, p = 0.677), nor significant interactions with the prompts (VERB × point in time: F (1,77) = 0.112, p = 0.739; STRAT × point in time: F (1,77) = 3.231, p = 0.076), or their combination (VERB × STRAT × point in time: F (1,77) = 0.003, p = 0.953).

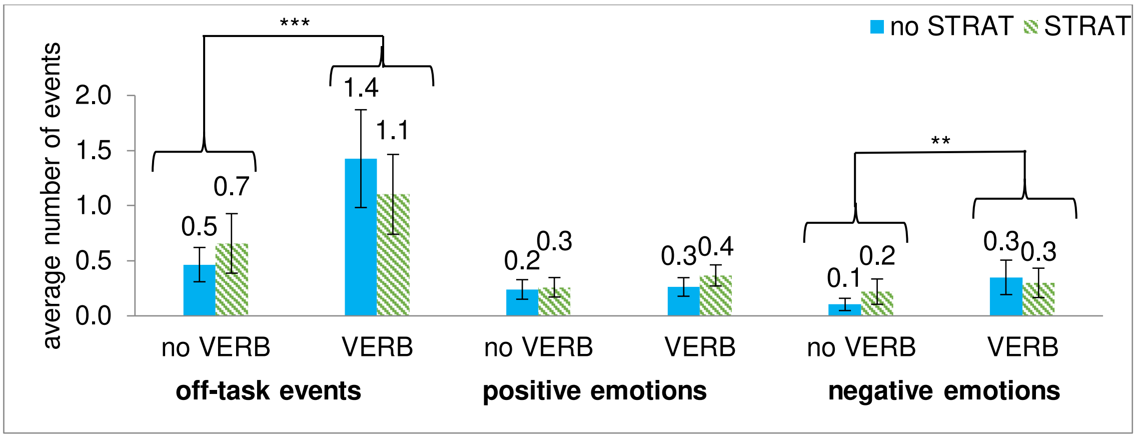


Figure 9. Results of study III as published in Schmitt and Weinberger (2019): Average number of off-task events, and positive and negative emotions per problem analyzed; main effect of VERB on off-task behavior and negative emotions. Error bars represent 95% CIs. *p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001.

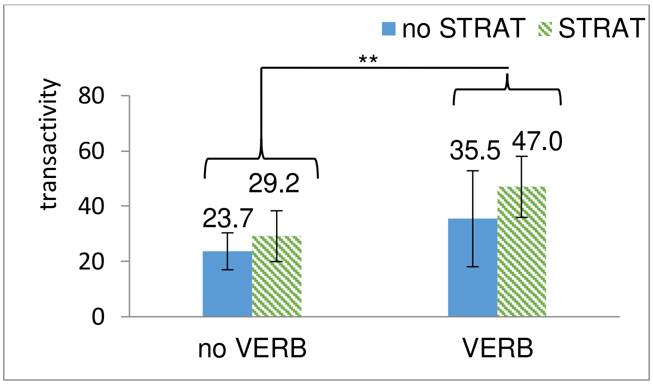


Figure 10. Results of study III as published in Schmitt and Weinberger (2019): Scores in transactivity; main effect of VERB on transactivity. Error bars represent 95% CIs. *p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001.

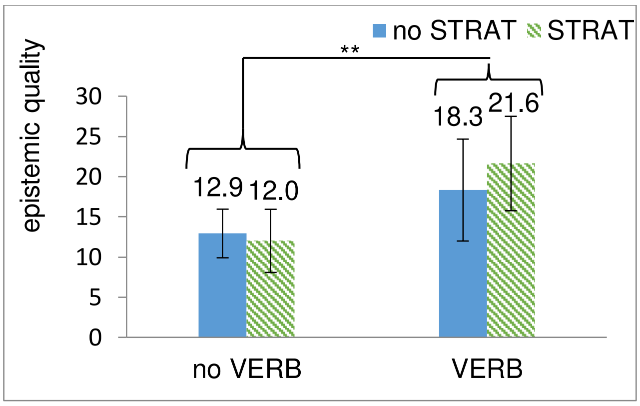


Figure 11. Results of study III as published in Schmitt and Weinberger (2019): Scores in epistemic quality; main effect of VERB on epistemic quality. Error bars represent 95% CIs. *p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001.

Table 16
Results of Study III as Published in Schmitt and Weinberger (2019): Observed Total Numbers, Means (SD) per Participant, and Frequency Percentages for Epistemic Quality and Transactivity Categories

Epistemic quality				Transactivity			
Category	Total number	M (*SD*)	Frequency percentage	Category	Total number	M (*SD*)	Frequency percentage
Off-topic utterance	43	.3 (.7)	1.4 %	Externalization	1494	10.1 (5.1)	42.8 %
Regulation of the interaction	2186	14.8 (8.4)	59.8 %	Externalization as reaction	86	.6 (.8)	2.5 %
Concrete task-related regulation	909	6.1 (3.5)	25.9 %	Acceptance	451	3.0 (2.7)	11.3 %
Abstract content-related regulation	168	1.1 (1.2)	4.3 %	Refusal	277	1.9 (1.5)	8.2 %
Procedural knowledge / strategies	99	.7 (.8)	3.2 %	Elicitation	715	4.8 (2.9)	20.0 %
				Integration	54	.4 (.6)	1.3 %
				Conflict-oriented consensus building	328	2.2 (2.2)	8.6 %
Not codable	159	1.1 (1.0)	5.4%	Not codable	159	1.1 (1.0)	5.4%

7.3.4. Discussion, limitations, and future work

Overall, the participants showed significant learning gains in near transfer tasks, but not in far transfer tasks. While the strategy prompts did not significantly influence any of the measured dependent variables, the verbalization prompts had versatile effects: They increased off-task behavior and negative emotions, but also improved dialogue quality, regarding both transactivity and epistemic quality, and did not affect learning gains.

The lack of effects of strategy prompts pointed at the consideration that this way of supportive scaffolding does not universally work, and that its effectiveness may be dependent on the type of learning environment. If the learning environment is already motivating and engaging, additional scaffolds may not be needed to keep learners on track. In support of this, Jackson et al. (1998) report that learners eventually turn off or simply ignore these sorts of prompts. Consequently, learners are unable to profit from prompts that they do not engage with. Future research may focus on developing domain-specific prompts targeting the specific misconceptions of the domain, in order to make their potential benefit more salient to the learners. The strategy prompts were integrated into *Proportion* in a fairly subtle way to increase their acceptance, however, one may also consider to intensify the strategy prompting, while taking concerns regarding over-scripting (Dillenbourg, 2002) into account. Beyond motivational factors, cognitive factors, such as the challenge to understand and apply the prompts, may account for the lack of effects; more time and practice may have helped in translating the prompts into productive behavior (Schoenfeld, 1987).

The verbalization prompts had multifaceted effects. The positive effects on discussion quality (transactivity and epistemic quality) are an encouraging finding as these qualities are considered a key for successful CSCL; this study showed that even primary school students' dialogue can be prompted successfully, which is in line with prior research on scaffolding young learners (Gelmini-Hornsby et al., 2011; Gijlers et al., 2013). However, verbalization prompts also raised off-task behavior and negative emotions. Also Chen and Law (2016) observed detrimental effects of question prompts on motivation in a tablet-based learning environment. Potentially, the prompts did not match with learners' subjective needs, as they imposed an additional difficult task, whose value may have not been apparent to the learners. Finally, the verbalization prompts did not foster learning beyond the overall positive learning gains (independent from condition). Due to the higher-quality discussion with verbalization prompts, a positive effect on learning gain could have been expected, however, the diminished task focus may have negatively compensated this effect, as sustained task focus is an important factor for learning (e.g., Cohen & Lotan, 1995). Further, other effects such as the *verbal overshadowing* effect (verbalization of hard-to-verbalize representations can corrupt performance; Schooler, 2002), or social effects (for young learners, collaborating with friends is more effective; van Dijk et al., 2014) may explain why learning gains were not specifically fostered through verbalization prompts.

The identified limitations of the study comprise the potentially limited generalizability of the results, the non-blind video coding, the data reduction, and the aggregation of individual categories of the dialogue quality variables. These limitations were taken care of by including a relatively diverse sample of participants, by ensuring sufficient coder training, by selecting representative data samples, and by supplying additional qualitative case analyses.

For both prompts, future research may focus on finding ways to make the prompts more personally relevant and meaningful to the learners in order to foster productive engagement with them. For example, strategy prompts may be designed to tackle the typical misconceptions of a domain, and verbalization prompts may benefit from a teacher introduction of how to productively handle them. Furthermore, it may be worthwhile to develop dynamically adapting prompts that sense learners' current needs and respond appropriately to them. Also, more work needs to be done regarding orchestration of embodied and reflective activities. For instance, rather than tightly integrating these activities as done in this study, homogeneous phases of immersive embodied learning vs. verbal reflection and abstraction phases may make sense, which leads to further open research questions, like the sequencing of these experiences (see e.g., Ottmar & Landy, 2017), and how to support translation between both experiences, and ultimately, to individual test performance.

7.4. Study IV: The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets[6]

Study IV is currently in preparation. The focus of this study is on investigating the effects of learning process heterogeneity in embodied collaborative learning on two dimension: Cognitive and bodily.

7.4.1. Theoretical background: Learning in collaborative embodied learning environments, embodied cognition for collaborative learning, collaborative learning and emotions, heterogeneity in collaborative learning, proportional reasoning and learning with Proportion

Drawing on theories of embodied cognition (e.g., Glenberg, 2010), and collaborative learning (e.g., Dillenbourg et al., 2009), hands-on learning environments with bodily activities (e.g., Reinholz et al., 2010) coupled with a collaborative learning setting may support children's

---

[6] Study IV is in preparation:
Schmitt, L. J., Tsovaltzi, D., & Weinberger, A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Revisions required in a review process may lead to differences between the summary provided in this thesis and the final article publication.

acquisition of challenging mathematical concepts like proportional reasoning (e.g., Mix et al., 1999). An embodied collaborative learning environment allows for rich interactions between learners (e.g., Davidsen & Ryberg, 2017), which afford an analysis of learners' processes on multiple modes. While it is known that learner heterogeneity in prior characteristics (e.g., Manske, Hecking, Chounta, Werneburg, & Hoppe, 2015), and the quality of learning processes (e.g., Webb, 1991) alike strongly influence learning, little is known about the role of heterogeneous vs. homogeneous learning processes, for example, on a cognitive and a bodily dimension.

In collaborative learning, respectively CSCL, a main research focus traditionally was on adults and their verbal interactions, for example, online argumentation (e.g., Tsovaltzi et al., 2015). However, also children's collaborative learning – computer-mediated or co-present (Schmitt & Weinberger, 2018 / study I, section 7.1), and the role of bodily processes (Lund et al., 2019) have lately received increased attention.

A fundamental theoretical assumption of study IV is embodied cognition (e.g., Niedenthal et al., 2005; Spackman & Yanchar, 2014; Wilson, 2002), i.e., the assumption that bodily experiences are central for information processing and reasoning. *Online embodiment* refers to embodied processing during live contact or interaction with a stimulus (e.g., Niedenthal et al., 2005; Wilson, 2002), and online embodiment may be integrated into so-called *embodied learning environments* through bodily activities. Touch interfaces are instances of environments for active and direct bodily manipulation (e.g., Black et al., 2012; Ottmar & Landy, 2017), and they can be beneficial for supporting co-present collaborative learning. For example, they can foster bodily awareness of the learning partners (e.g., Jakobsen & Hornbæk, 2016), and cognitive processes (e.g., Mercier & Higgins, 2013). In addition to adding embodied learning experiences to more traditional ones in a sequence of activities (Black et al., 2012), one may also integrate both approaches, i.e., bodily involvement tightly coupled with verbal reflection activities (see Schmitt & Weinberger, 2019 / study III, section 7.3), as verbal and bodily activities together and in interaction are central for forming a good understanding of a subject (e.g., Gomoll, Hmelo-Silver, Tolar, Šabanovic, & Francisco, 2017; Rick, 2012b). Scaffolding embodied learning experiences and providing explicit opportunities for reflection on underlying concepts may be required (e.g., Schneps et al., 2014), however, can result in mixed effects as shown by study III (Schmitt & Weinberger, 2019; section 7.3). Consequently, there is a need for more research on the interplay of bodily and cognitive learning processes, for example, on the relation between bodily expression of emotions and the epistemic quality in discourse, to advance our understanding of embodied collaborative learning.

Emotions play an important role for collaborative learning processes, for example, the intensity of emotions relates to the quality of talk (e.g., Polo et al., 2016), and maintaining a productive emotional climate is crucial for leaners' discussions (Isohätälä et al., 2018). The body is involved in emotional experiences, both in shaping them as well as in expressing them (Ekman, 1992; Strack et al., 1988), thus, emotional processing is embodied (Niedenthal et al., 2005). Collaborative learning creates a situation that affords bodily involvement, such as bodily expression of emotions (Isohätälä et al., 2018), or bodily movements for explaining concepts or physically regulating interactions (Davidsen & Ryberg, 2017).

Heterogeneity of collaborative group members can influence learning-relevant variables, for example, heterogeneous domain knowledge and motivation was positive for performance (Manske et al., 2015). Often, heterogeneous group formation based on cognitive variables (e.g., level of prior knowledge) was found to be advantageous for learning, which can be explained by the emergence of productive cognitive conflicts (Jorczak, 2011). However, heterogeneity in prior knowledge can also cause dominant behavior (ter Vrugte et al., 2015). Moreover, heterogeneous group formation (in this case, based on the individual goal structures) can also positively affect emotional appraisals of collaborative learning (Wosnitza & Volet, 2012). There is a multitude of possible variables for group formation, such as social preference (van Dijk et al., 2014), or productivity (Hamilton, Nickerson, & Owan, 2003), with mixed results, depending on the chosen variable. Also, different learners may profit from heterogeneity differently (Cheng et al., 2008), and the quality of the group processes (Cheng et al., 2008), and the learning environment design (Asterhan et al., 2013) can also change the effects of different ways of group formation. Therefore, it is difficult to come to clear conclusions on the effects of heterogeneous vs. homogenous group formation. It may be worthwhile to also address the role of heterogeneous vs. homogeneous learning processes – and not just prior characteristics, like prior knowledge – during a collaborative learning phase, which is a largely unexplored research topic. The few available findings indicate that heterogeneity in bodily expression of emotions can be beneficial for collaboration (Rick et al., 2015), possibly through an increased awareness of the ongoing bodily processes. This awareness may lead to deliberate attempts to share individual embodied representations (see Barsalou, 1999), and be turned into productive cognitive and bodily interactions, but may also be distracting and require unproductive effort for regulating each other. Heterogeneity in cognitive processes similarly may either enhance cognitive processes and learning (e.g., Curşeu, Schruijer, & Boroş, 2012; Rosebery, Ogonowski, DiSchino, & Warren, 2010), or negatively affect group climate and lead to process losses (Curşeu et al., 2012). In sum, positive as well as negative effects of process heterogeneity can be expected.

Proportional reasoning (e.g., Boyer et al., 2008) is challenging for children, but may benefit from embodied learning experiences (e.g., Reinholz et al., 2010). With the *Proportion* learning environment (e.g., Rick, 2012b), young learners can collaboratively and directly manipulate two bars to represent proportional relations. The collaborative embodied scenario fosters a high bodily awareness of each other's bodily activities (Rick, 2012b), talk, and emotional expressions. Proportion is assumed to target individual embodied processing, and to allow for free bodily interactions and expression of emotions, for an embedded and embodied social experience.

This study explores the research question of the role of bodily and cognitive learning processes, and to what extent heterogeneous or homogeneous processes influence learning in a collaborative embodied scenario. As prior research on these aspects is sparse, the hypotheses are tentative and non-directional. Summarized, the study tests the effects of heterogeneity in bodily processes on the quality of cognitive processes (hypothesis 1), the effects of heterogeneity in bodily processes on the groups' performance (hypothesis 2), and the effects of heterogeneity in cognitive processes on the groups' performance (hypothesis 3).

7.4.2. Methods

The sample comprised n = 80 participants, the sub-sample receiving verbalization prompts from Schmitt and Weinberger (2019) / study III. The participants were, on average, 10 years old (SD = 6.9 months), with 51.3 % males. Dyads were formed randomly. Regarding the experimental procedure, see also section 7.3.2, students learned collaboratively with *Proportion* for 40 minutes, and before and after the learning phase, responded to a survey and a math test individually.

Coding of the multimodal learning process variables is reported in Schmitt and Weinberger (2019) / study III. The analyzed process variables comprise the bodily expression of emotions, off-task behavior, and the quality of talk. These process variables were video coded with sufficient inter-rater reliability, specifically, values between $.77 \leq \alpha \leq .92$ (Krippendorff's alpha; Krippendorff, 2012), see Schmitt and Weinberger (2019). In addition, results of the math post-test are presented to assess knowledge outcomes and convergence. Heterogeneity was determined through the *Coefficient of Variation* (*CV*; Weinberger, Stegmann, & Fischer, 2007). Using a median split, the sample was divided into heterogeneous (upper half of CV) vs. homogeneous (lower half of CV) groups.

The three main groups of dependent variables – *bodily processes*, *cognitive processes*, *performance* – consisted of the following individual variables: Bodily processes reflect the bodily expression of emotions (both positive and negative emotions). Cognitive processes encompass learners' observable task focus, as well as the degree of transactivity and epistemic quality in their dialogue. Finally, learners' performance comprised the knowledge outcomes, measured through the math test, knowledge convergence (CV in math test within groups), and efficiency, i.e., the amount of problems learners solve in the app.

The statistical analysis is based on MANOVAs for testing the three hypotheses, a mediation analysis based on the "bootstrap confidence interval for the indirect effect" approach (Hayes & Rockwood, 2017, p. 44) to further explain the outcomes of hypotheses 1 and 2, and a qualitative case analysis to illustrate the quantitative findings.

7.4.3. Results

Regarding general learning gains, the participants showed, overall, significant improvements from pre- to post-test in near transfer tasks (t (39) = -2.980, p = .005, d = .257), but not in far transfer tasks (p > .05).

Next, the MANOVA results for the three main hypotheses are presented. Regarding hypothesis 1, there was an overall large significant effect of heterogeneity in bodily processes on cognitive processes: F (3, 26) = 4.732, p = .009, *Pillai's Trace* = .353, η_p^2 = .353. Specifically, a large negative effect on transactivity was marginally significant with F (1, 28) = 3.796, p = .061, η_p^2 = .119, i.e., homogeneous bodily processes fostered, by trend, transactivity, which is in support of the first hypothesis. Heterogeneity of bodily processes did not significantly affect epistemic quality or task focus (p > .05). Regarding hypothesis 2, there was an overall large significant effect of heterogeneity in bodily processes on performance: F (3, 31) = 3.858, p = .019, *Pillai's Trace* = .272, η_p^2 = .272. Specifically, a large positive effect on efficiency was significant with F (1, 33) = 9.179, p = .005, η_p^2 = .218, i.e., heterogeneous bodily processes fostered the number of solved problems, which is in support for the second hypothesis. Heterogeneity of bodily processes did not significantly affect knowledge outcomes or knowledge convergence (p > .05). Regarding hypothesis 3, there was an overall large significant effect of heterogeneity in epistemic quality on performance: F (3, 30) = 3.096, p = .042, *Pillai's Trace* = .236, η_p^2 = .236. Specifically, a medium negative effect on knowledge outcomes just missed significance with F (1, 32) = 4.025, p = .053, η_p^2 = .112, and a large negative effect on knowledge convergence was significant with F (1, 32) = 8.120, p = .008, η_p^2 = .202, i.e., homogeneous epistemic processes fostered knowledge convergence, and by trend, knowledge outcomes, which is in support for the third hypothesis. There were no further significant effects of heterogeneity in epistemic quality on efficiency, or of heterogeneity in the other two cognitive variables on performance ($p > .05$).

The results regarding hypotheses 1 and 2 led us to formulate a mediation hypothesis stating that transactivity may be a mediator for the effect of heterogeneity in bodily processes on efficiency. This mediation model (see Figure 12) was supported by the statistical analysis (Hayes & Rockwood, 2017): The bootstrapped (5000 samples) unstandardized indirect effect of X (heterogeneity of bodily processes) on Y (efficiency) equaled -8.242, and the 95% confidence interval ranged from -17.452 (BootLLCI) to -.5160 (BootULCI). I.e., transactivity does serve as a mediating variable for the effect of heterogeneity in bodily processes on efficiency.

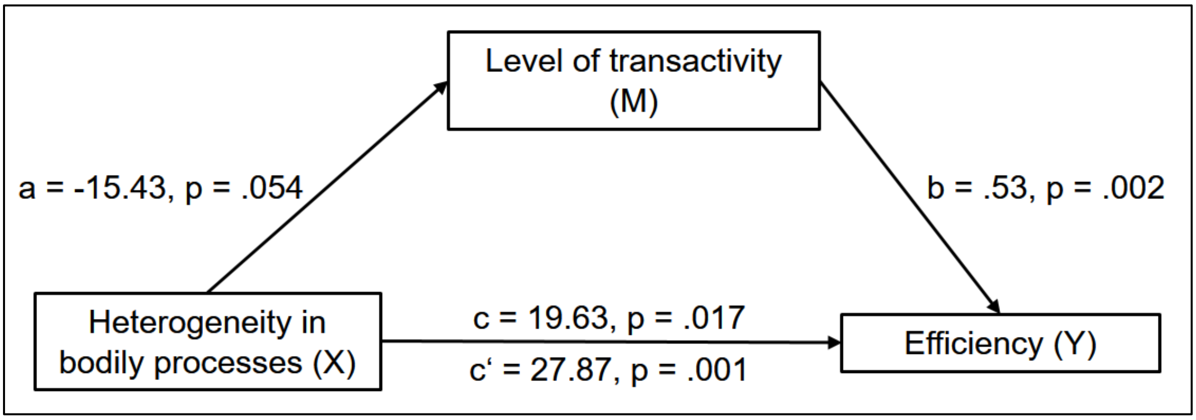


Figure 12. Mediation model presented in Schmitt, Tsovaltzi, and Weinberger (in preparation) / study IV: Transactivity as a mediator for the effect of heterogeneity in bodily processes on efficiency.

Further, study IV provides a case analysis of the problem solving processes of a bodily heterogeneous dyad, Willy and Fiona. Summarized, this dyad strongly relies on bodily interactions rather than verbal discussions. While they are not very transactive in their dialogue (transactivity value of 23.50 compared to the sample's mean of $M = 42.3$ ($SD = 28.37$)), they display a high bodily efficiency with 128 solved problems (compared to the sample's mean of $M = 93.75$ ($SD = 23.78$)). They use the fingers to measure proportions, to guide the partner (pointing), and to engage in and resolve physical conflict. Thus, the case analysis illustrates how a bodily heterogeneous dyads approaches the task in an embodied way. See Figure 13 and Table 17 for details of their ongoing verbal and bodily processes.

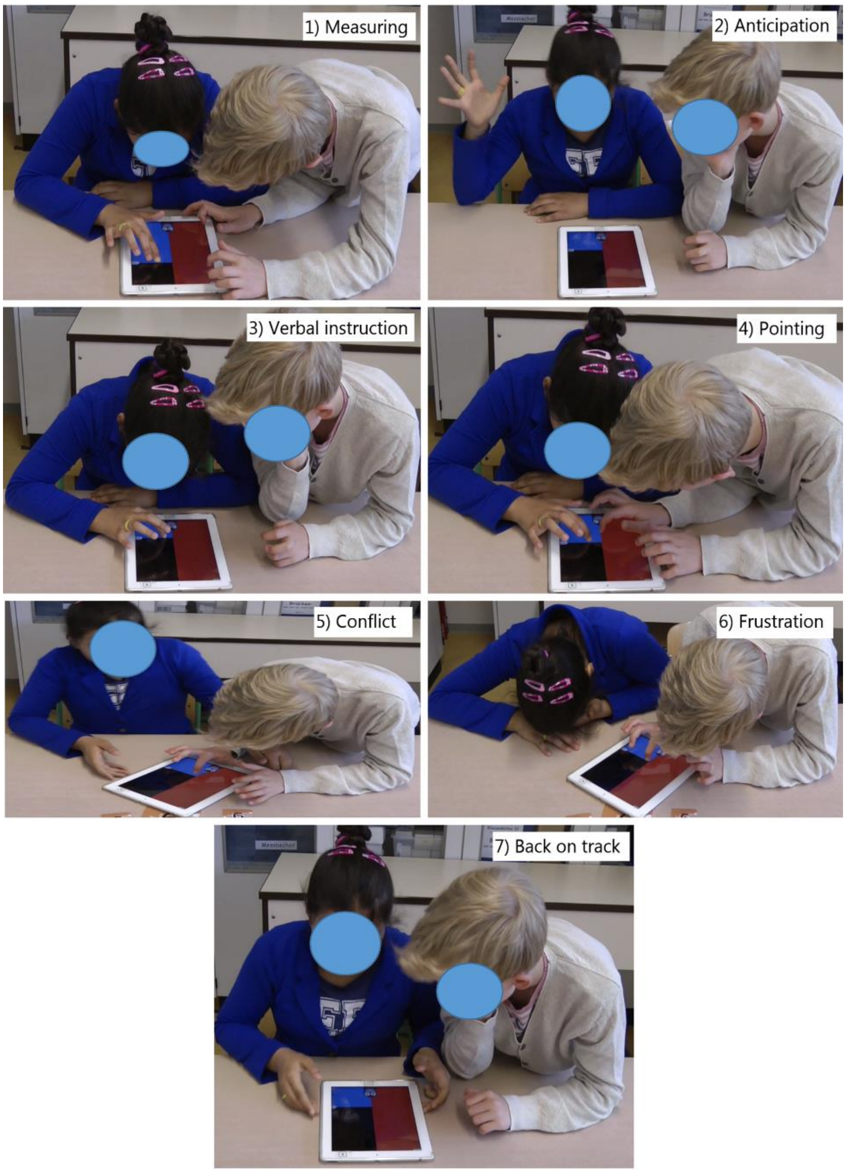


Figure 13. Screenshots of the task progress of two bodily heterogeneous learners during the solution of a 'Proportion' task; figure from Schmitt, Tsovaltzi, and Weinberger (in preparation) / study IV.

Table 17
The Unfolding Verbal and Bodily Processes of Two Bodily Heterogeneous Learners During the Solution of a 'Proportion' Task; Table from Schmitt, Tsovaltzi, and Weinberger (in preparation) / study IV

Time	Verbal interaction	Bodily interaction	Comment
19:50 – 19:52		Willy: Playing with the orange bar Fiona: Playing with the blue bar	No clear goal or solution strategy recognizable
19:53 – 19:56	Willy: *Hey, what are you doing?* Fiona: *I am looking for the 10.*		Starting point for goal-directed behavior
19:57 – 20:00		Willy: Measuring with his fingers Fiona: Slowly moves the blue bar down	Embodied approach to task solution (measuring, manipulating) [screenshot 1]
20:01	Willy: *To this line here.*	Willy: Points at a position on the tablet that Fiona should move to	Willy guides Fiona verbally and bodily
20:01 – 20:02		Willy: Keeps on pointing with his fingers Fiona: Moves the blue bar to the position that Willy points at	Bodily instruction through pointing gesture
20:03	Willy: *Stop!*		Verbal co-regulation
20:04 – 20:05		Fiona: Touches the owl and dramatically throws hands in the air	Fiona shows bodily anticipation of success [screenshot 2]
20:06	Solution is not yet correct; owl gives feedback: "almost"		
20:06	Willy: *More down.* Fiona: *Oh!*		Willy guides Fiona verbally [screenshot 3]
20:06 – 20:08		Fiona: Slowly moves the blue bar down	Fiona appropriates the task bodily

Time	Verbal interaction	Bodily interaction	Comment
20:09	Willy: *No!*	Willy: Tries to wipe away Fiona's hand	Willy guides Fiona bodily and verbally
20:10 – 20:12	Willy: *More up…no… you are too much up… there!*	Fiona: Moves the blue bar up	Willy guides Fiona verbally
20:13 – 20:17	Willy: *Exactly on here!*	Willy: Points at a position on the tablet that Fiona should move to Fiona: Fiona: Slowly moves the blue bar down	Willy guides Fiona bodily and verbally and Fiona bodily executes the instructions [screenshot 4]
20:18	Fiona: [Mumbles something]	Fiona: Touches the owl	
20:18	Solution is still not correct; owl gives feedback: "almost"		
20:19 – 20:24	Willy: *Let me try it!*	Willy: Pulls the tablet to himself; physically dominates access to the tablet; carefully moves the blue bar up Fiona: First tries to get access back, then puts the head down on the table in frustration	A little conflict regarding physical access emerges between the learners; Fiona bodily disengages the task and seems to express despair [screenshot 5 and 6]
20:24	The problem is solved		
20:24	Willy: *So!*		
20:25		Willy: Puts the tablet back in the middle Fiona: Physically joins back in	The physical conflict is resolved through physical action (giving back access) [screenshot 7]

7.4.4. Discussion, conclusions, limitations, and future research

Overall, the study shows that heterogeneous as well as homogeneous processes can have positive effects – depending on the dimension of heterogeneity. Homogeneous bodily processes, by trend, furthered transactivity, and homogeneous cognitive processes increased knowledge convergence, and by trend, knowledge outcomes. In contrast, heterogeneous bodily processes increased bodily efficiency, i.e., the number of solved problems in the app. Transactivity mediates the relationship between heterogeneity in bodily processes and efficiency, and this mechanism was further substantiated by the provided case analysis of one bodily heterogeneous dyad.

The tentatively positive effect of homogenous bodily processes on transactivity (hypothesis 1) may be due to the decreased effort of co-regulating the rather low amount of bodily divergences, therefore, having more capacity for engaging in collaborative cognitive processing. In contrast, bodily heterogeneous learners did not seem to be able to resolve the bodily divergence through higher cognitive processes; possibly, the emotional group climate in these cases was not conducive for their discussion quality (see Isohätälä et al., 2018). The verbalization prompts for reflection may be readily adopted by bodily homogeneous learners and result in higher quality talk, but rather interfere with the body-based interaction styles of bodily heterogeneous learners. These bodily heterogeneous learners seem to rather focus on their bodily interactions and solve many problems in the app (hypothesis 2), possibly because the bodily aspects of their interaction and the task become more salient, affording embodiment (see Barsalou, 1999). The bodily success in the app, however, did not transfer to knowledge outcomes or convergence, thus, may benefit from added support to ease the transfer towards abstraction of the embodied concepts. The positive effects of homogenous cognitive processes on knowledge outcomes (by trend) and knowledge convergence (hypothesis 3) may be due to a high understanding of each other's contributions, i.e., the learners share a similar way of expressing their understanding and can make sense of each other more easily. The homogeneity may also increase the learners' readiness to accept their partner's contributions (Marks, Copland, Loh, Sunstein, & Sharot, 2019). In contrast, cognitively heterogeneous groups may have been unable to cope with their cognitive divergences combined with the required bodily interactions in the app, or may have suffered from patterns of dominance (ter Vrugte et al., 2015).

According to the mediation analysis, transactivity is a mediator for the effect of heterogeneity in bodily processes on efficiency. Depending on their bodily heterogeneity, dyads adopt different interaction styles: Bodily homogeneous dyads focus on high-quality dialogue, and therefore, less on bodily activities and efficiency ("the talkers"); bodily heterogeneous dyads do not engage much in high-quality dialogue, but focus on their bodily progress and solve many problems in the app ("the makers"). While there is an interdependency between bodily and cognitive processes, with different foci dependent on the group's bodily heterogeneity, these two interaction styles do not differently affect learning outcomes. The provided case study of a bodily heterogeneous dyad (Willy and Fiona) illustrates this finding by setting out their focus on bodily interactions.

To conclude, embodied collaborative learning can help children acquire proportional reasoning skills, and the level of heterogeneity in their ongoing cognitive and bodily processes shapes their cognitive interactions and affects their performance. These findings underline the relevance of an embodied perspective for designing and analyzing learning environments, and for taking into account group formation aspects as a relevant factor for learners' interactions. Additional prompting may not only focus on verbal abstraction (as in Schmitt & Weinberger, 2019 / study III), but potentially, also target learners' process heterogeneities, along with support of how to utilize them for their learning, for example, to align their way of expressing their understanding.

The limitations of this study comprise the explorative nature of the research, and the rather small sample size. Future research may experimentally vary group process heterogeneity with a larger sample, and try to generalize the found effects to other contexts. The results of the present study can advance our theoretical understanding of the role of multimodal (heterogeneous) learning processes, which is a crucial prerequisite for the development of automatic process analysis and group formation algorithms (e.g., Erkens et al., 2016; Gweon et al., 2013).

8. General Discussion and Outlook

This concluding chapter will summarize the research presented in this thesis (section 8.1), comment on strength and limitations of this research (section 8.2), provide directions for future research (section 8.3), and finish with conclusions drawn from this work (section 8.4).

8.1. Summary of the Research of the Present Thesis

In the present thesis, the focus was on exploring children's co-present collaborative and embodied learning with a touch interface. Upon introducing this sort of collaborative learning scenario in chapters 1 and 2, the following chapter 3 elaborated on theories of embodied cognition and learning, and the integration of bodily activities into learning environments. In the subsequent chapter 4, an overview on learning-relevant process variables was provided. Chapter 5 introduced the learning domain – proportional thinking – and the *Proportion* learning environment, and chapter 6 elaborated on research methods to capture children's experiences and learning processes on multiple dimensions.

In the four studies which were presented next, the focus was first on reviewing and discussing computer-supported learning opportunities for young learners (Schmitt & Weinberger, 2018 / study I, section 7.1). Computer support may come in various forms, for example, with desktop computers, handhelds, or tabletops, and can facilitate computer-mediated as well as co-present learning scenarios. Study I also commented on typical problems of these scenarios, such as lack of participation or questions of how learners regulate and share access to a jointly used device, and finally reviewed solutions to improve collaborative learning opportunities, such as structuring interactions with scripts. Study II (Schmitt et al., 2019; section 7.2) focused on the challenge of adequately surveying children's technology evaluation and, to that end, reviewed existing instruments with their strengths and weaknesses, and motivated the need to develop a new one: BiCo. The BiCo instrument was then put to the test in a study comparing it with the well-established Smileyometer instrument. The results supported BiCo's utility as an instrument to be used in child computer interaction research. Study III (Schmitt & Weinberger, 2019; section 7.3) reports on a study which experimentally varied two sorts of prompts to support learners' strategic behavior, and to support high quality explanations and joint reflection while interacting in an embodied learning environment. The results showed that, while there were overall learning gains independent from the experimental condition, the strategy prompts did not have significant effects on learning processes and outcomes, however, the verbalization prompts led to more instances of off-task behavior and negative emotions, but also enhanced the quality of children's dialogue. Finally, to better understand learners' interactions and interdependencies between different dimensions of a learning process, study IV (Schmitt et al., in preparation; section 7.4) focused on the influence of learning process heterogeneity, specifically heterogeneity in bodily and cognitive processes, on the groups' cognitive processes and performance. The results showed that both heterogeneous and homogeneous processes can be beneficial – depending on the dimension of interest. For example, heterogeneous bodily processes increased efficiency, but homogeneous cognitive processes increased knowledge convergence. Further, transactivity mediated the relationship between heterogeneity in bodily processes and efficiency.

With the *Proportion* environment, children can acquire skills in mathematics by collaboratively interacting with a joint touch device, including bodily involvement and movements. Drawing on embodied cognition for designing learning environments does not necessarily imply that children are entirely freely moving in the whole classroom or even beyond. Embodied learning may also be realized on a smaller scale with touch tablets that afford bodily movements to interact with the representations displayed on the screen. This sort of learning with a tablet app may be fairly easily integrated into a regular school lesson, i.e., without changing the whole learning setting entirely. For example, a teacher who wants to introduce proportional reasoning to their class may start with one lesson supported by the tablet app, where learners can collaboratively explore and 'grasp' concepts of proportionality, and in a following lesson, the teacher may then proceed to facilitate joint reflection on these experiences and start to introduce formal rules and symbols. Such a way of arranging these learning phases would allow for both: A certain amount of bodily involvement, as well as the traditionally more abstract and formal approach to the topic (which is important as well). Taken together, this may be a powerful approach for fostering a more holistic understanding of a domain, without radically changing teaching practices that will likely continue to be somewhat formal and inside classrooms. This approach may not only be possible with *Proportion*, but may be transferred to other domains and learning environments as well.

8.2. Strengths and Limitations of the Present Thesis

One major strength of this thesis is the multimodal data analysis of multiple dimensions of learners' interactions. For example, in study III (Schmitt & Weinberger, 2019; section 7.3), a rich data set (video data of n = 162 participants) served as the basis to measure learners' processes on a verbal dimension with a focus on epistemic quality and the degree of transactivity, and on a non-verbal dimension including learners' emotions and task focus. The multifaceted effects of the verbalization prompts on the learning processes would not have been detected without this sort of multimodal process data. The insights generated by this study support the development of theories about collaborative embodied learning, can inform the design of prompts and learning environments, and also inspired further research questions and the follow-up analysis in study IV (Schmitt et al., in preparation; section 7.4).

Further, the experimental studies (studies II, III, and IV) took place in a natural setting, i.e., inside the classrooms of the schools and during normal school time, while still adhering to experimental rigor. To illustrate, it was made sure that the children were assigned to groups and conditions randomly, and the instructions were standardized. Like this, typical biases which are unavoidable in a – highly controlled but artificial – lab setting, or a – more realistic but less controllable – field study situation, were mitigated. The experimental approach allowed, for example, to systematically vary experimental conditions, in order to test hypotheses and draw conclusions on the effects of different learning environment designs (Schmitt & Weinberger, 2019 / study III, section 7.3), or generated quantitative data to statistically compare two surveying instruments for children (Schmitt et al., 2019 / study II, section 7.2).

One limitation of this thesis is the assessment of proportional reasoning skills through a (pen and paper) math test. While assessing knowledge with a test is common practice and has a lot of advantages, for instance, efficiency and the suitability for statistical analysis, this form of assessment may be less ideal for capturing the sort of learning that happens in an embodied learning environment, and that may be more bodily and intuitive in nature. For example, in study III (Schmitt & Weinberger, 2019; section 7.3), there were significant learning gains in near transfer tasks, which, however, did not vary as a function of the experimental condition. While the verbalization prompt factor affected several process variables significantly (e.g., higher dialogue quality), the learners may have been unable to transfer these sort of skills and knowledge gained in an interactive embodied collaborative setting onto an individually administered pen and paper math test (see also, Schmitt & Weinberger, 2019 / study III, section 7.3). Potentially, a more sensitive approach of testing learners' embodied understanding would have made sense, for example, by means of another collaborative task based on bodily interactions. Gestures themselves can serve as an indicator for mental representations (Alibali, Bassok, Solomon, Syc, & Goldin-Meadow, 1999), and it seems plausible that an understanding which has been gained in a collaborative embodied learning setting is difficult to be expressed outside of that context. However, individual outcome assessment does have a pivotal role in assessing learners' ability to transfer from the collaborative to the individual mode (Kuhn & Crowell, 2011). A possible solution for this methodological challenge could be to combine the strengths of both approaches and include both forms of assessment together in a study – collaborative and based on bodily activities, plus individual formal testing.

8.3. Beyond this Thesis: Directions for Future Research

Future research may build up on the research presented in this thesis by following up on the discussion regarding collaborative learning in an embodied learning environment. How can we measure learning processes on multiple dimensions? To what extent do individual learning processes relate to other learning processes and to knowledge acquisition? How can learners be further supported in productively interacting with a device and their learning partner? To what extent do learning processes change when learners with similar or different ways of working are grouped together? More concretely, as proposed in the four studies of this thesis, future research may address the following questions and research directions:

Study I (section 7.1) calls for further developing instructional means for supporting effective CSCL in school settings, for instance, through awareness tools, or scripting / scaffolding, or a combination of both (Schmitt & Weinberger, 2018). Further, there is a need to continue investigating orchestration of individual and collaborative learning phases, and to supplement these phases with appropriate technologies and support (e.g., scaffolding) in order to enable the learner to encounter diverse challenges and develop a robust understanding (Schmitt & Weinberger, 2018). Essentially, more work needs to be done in finding ways to effectively support learners in navigating different learning activities, being formal or informal, collaborative or individual, with (diverse) computer-support or without, and integrate these diverse experiences for enhanced learning (Schmitt & Weinberger, 2018).

Study II (section 7.2) points out the role of continuing to refine survey methods for research with children in the context of computer-supported learning. For instance, there is a need for having reliable survey instruments for a wide variety of target groups (children of different ages and background), or for assessing learning processes in diverse learning settings (Schmitt et al., 2019). Moreover, future research will need to investigate children's perceptions of interacting with different survey instruments, to find out, for example, which surveying methods are more or less time-consuming or taxing, to what extent an instrument should be labeled, or how to ensure good item wording (e.g., two bipolar statements) (Schmitt et al., 2019).

Study III (section 7.3) demonstrates the need to continue research on the content, timing, and intensity of scaffolding prompts. For example, different phases of a learning activity may benefit from tailored support to address learners' misconceptions rather than providing overall strategic support and requesting verbalizations (Schmitt & Weinberger, 2019). Future research may focus on designing prompts in a way that they are sufficiently noticed and adopted by the learners, without overly interfering with their natural learning processes and potentially disengaging the learners. A teacher may have a central role in this process by making the role of the prompts salient to the learners and support them in using them effectively. A technological solution to this challenge would be to make the scaffolding adaptive with a computer system gathering the relevant learner information and providing the sort of support that learners need in the right moment in time (Schmitt & Weinberger, 2019). Last, hands-on bodily activities on one hand, and verbal reflection and abstraction on the other hand, seem to be two different worlds. While there is the potential to combine their strengths and make learning more productive by enabling both – embodied experiences and abstraction through verbalization – it is a matter of future research of how to sensibly orchestrate these two worlds (Schmitt & Weinberger, 2019). For example, rather than integrating both approaches into one learning activity, it may make sense to experience them sequentially. Future research will need to follow up on these ideas, for instance, by investigating how to support the learners in drawing connections between hands-on vs. verbal abstraction learning phases (Schmitt & Weinberger, 2019).

Study IV (section 7.4) highlight the relevance of heterogeneous vs. homogeneous processes, and not only prior characteristics, for learners' interactions and learning achievements, and therefore, also identifies the need for further research and refinement of theories regarding this rather unexplored research topic (Schmitt et al., in preparation). Moreover, this study calls for developing scaffolding methods, which go beyond focusing on abstraction or strategic behavior (as in Schmitt & Weinberger, 2019 / study III), but also provide support for turning heterogeneous learning processes into productive interactions. For example, to prompt cognitively heterogeneous learners to engage in a common language and way of talking about their individual understanding, or to specifically prompt bodily heterogeneous learners to also focus on high quality dialogue, and verbally expressing their embodied understanding in addition to their bodily activities, in order to accomplish the transfer towards abstraction of embodied concepts (Schmitt et al., in preparation).

In addition to the directions for future research which have been raised in the four studies as presented above, the concept of transactivity may be explored further. As it was conceptualized in the studies of this thesis and in many others, transactivity is a quality expressed through verbal contributions, i.e., learners referring to their learning partner in their written or oral contributions. However, with regards to co-present collaborative learning with a shared device, the transactivity concept may need to be expanded. Learners may not only verbally, but also behaviorally build up on their learning partner's activities and utterances (see also, Marshall, Hornecker, Morris, Dalton, & Rogers, 2008; Roschelle & Teasley, 1995). For example, one learner may start to physically manipulate an object or representation, but not finish the activity, and the second learner may come into play and pick up the move and complement and finish it. Or one learner may perform a bodily activity successfully, which is then copied by the second learner who imitates said bodily activity, without any verbal reference to this behavior. An analysis into this line of thought is being performed by Davidsen and Ryberg (2017), who, for instance, observe the following: "(…) he now moves his right hand close to Nathalie's hand and shepherd her hand to the correct position through touch and movement. However, Peter only shows Nathalie the right place of the vertical line through his shepherding movement and while he retracts his hand Nathalie moves her hand to the right and touches the vertical line and then moves it to the correct place" (p. 80). This episode could be conceptualized as a transactive move, which would remain undetected if we only look at *verbal* transactivity. As can be seen in the transactivity coding scheme developed in the course of this thesis (see Table 9), bodily activities are being acknowledged to some extent already. For instance, in the *refusal* category, the rejection may be targeted towards a statement or *action*. While analyzing transactivity through verbal contributions is probably sufficiently in text-based computer-mediated learning scenarios, co-present and embodied scenarios may benefit from an extension of the transactivity concept towards also including bodily aspects. Research questions related to *bodily transactivity* may focus on: What is the role of bodily transactivity for learning? Can bodily transactivity be fostered through the learning environment design? What does the interplay between bodily and verbal transactivity look like?

Moreover, future research may take into account how learners' bodily processes develop over time, in order to learn more about how movements and gestures develop during a learning activity and shape collaborative processes (Davidsen & Ryberg, 2017). To that end, refining and automating multimodal interaction analysis methods may help (e.g., Davidsen & Ryberg, 2017). Research in the last years has made progress in developing automated analysis of processes, for instance, towards utilizing machine learning algorithms for automatically measuring transactivity (Gweon, et al., 2013), or classifying facial expressions of emotions from video data automatically (Harley, Bouchet, Hussain, Azevedo, & Calvo, 2015). This sort of real time assessment may generate new insights on the interactions over time, but also bears the potential for reacting to the learning processes as they unfold (Gweon et al., 2013), for example with the opportunity to group and re-group students based on their ongoing learning processes (see also, Tsovaltzi et al., 2019), or by generating relevant information for adaptive prompting.

Last, future research may focus on the generalizability of the present findings. To what extent are these results specific to the chosen context of fourth graders working collaboratively with touch screens in the mathematics domain? For example, recent research suggests differences in tangible vs. graphical interfaces in the domain of children's learning of programming (Sapounidis, Demetriadis, Papadopoulos, & Stamovlasis, 2019): While tangible interfaces were overall perceived as offering more opportunities for collaboration, this effect was also influenced by gender and age of the children, for instance, the effect was not present for younger children (aged 8-9 years), but for older children (aged 12-13 years) (Sapounidis et al., 2019). Widening the scope may include different technologies which also allow for bodily involvement, for instance augmented / virtual reality technologies, or to include learners from different age groups (e.g., adults) and in different domains (e.g., foreign language acquisition).

8.4. Conclusions

The present thesis adds to the knowledge about co-present collaborative embodied learning of children. The studies addressed research gaps in the field, such as providing an overview on CSCL in primary and secondary education (Schmitt & Weinberger, 2018 / study I, section 7.1), the development of novel methods to adequately survey children (Schmitt et al., 2019 / study II, section 7.2), addressing the combination of bodily hands-on learning with reflective abstraction and strategic behavior (Schmitt & Weinberger, 2019 / study III, section 7.3), and investigating the role of heterogeneous learning processes (Schmitt et al., in preparation / study IV, section 7.4).

In conclusion, learning collaboratively and with bodily involvement can be productive, but the effects are dependent on multiple conditions, for example, while verbalization prompts can improve learners' dialogue, they can come with side-effects like increased off-task behavior (Schmitt & Weinberger, 2019 / study III, section 7.3). On top of that, also the group formation may impact learning, for instance, bodily heterogeneous groups may lack verbal transactivity, but excel in their ability to solve the body-based tasks in the learning environment (Schmitt et al., in preparation / study IV, section 7.4). For the further development of theories of learning in embodied and collaborative contexts, it is crucial to continue this work of untangling these various determining factors underlying co-present collaborative embodied learning.

References

Abrahamson, D. (2017). Embodiment and mathematics learning. In K. Peppler (Ed.), *The SAGE encyclopedia of out-of-school learning* (Vol. 1, pp. 248–252). Thousand Oaks, CA: SAGE Publications Inc. doi:10.4135/9781483385198.n98

Alibali, M. W., Bassok, M., Solomon, K. O., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science*, *10*(4), 327–333. doi:10.1111/1467-9280.00163

Asterhan, C. S. C., Schwarz, B. B., & Cohen-Eliyahu, N. (2013). Conceptual change in proportional reasoning: Effects of collaboration, own / partner reasoning level and hypothesis testing. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Cooperative Minds: Social Interaction and Group Dynamics. Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 1780–1785). Austin, TX: Cognitive Science Society.

Baker, M., & Lund, K. (1997). Promoting reflective interactions in a CSCL environment. *Journal of Computer Assisted Learning, 13*, 175–193. doi:10.1046/j.1365-2729.1997.00019.x

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategical, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182. doi:10.1037//0022-3514.51.6.1173

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577–660. doi:10.1017/S0140525X99002149

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. doi:10.1146/annurev.psych.59.103006.093639

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. H. (2006). Common ground, complex problems and decision making. *Group Decision and Negotiation*, *15*(6), 529–556. doi:10.1007/s10726-006-9030-1

Bell, A. (2007). Designing and testing questionnaires for children. *Journal of Research in Nursing*, *12*(5), 461–469. doi:10.1177/1744987107079616

Belland, B. R., Kim, C., & Hannafin, M. J. (2013). A framework for designing scaffolds that improve motivation and cognition. *Educational Psychologist*, *48*(4), 243–270. doi:10.1080/00461520.2013.838920

Berkowitz, M. W., & Gibbs, J. C. (1983). Measuring the developmental features of moral discussion. *Merrill-Palmer Quarterly*, *29*(4), 399–410. Retrieved from https://www.jstor.org/stable/23086309

Black, J. B., Segal, A., Vitale, J., & Fadjo, C. (2012). Embodied cognition and learning environment design. In D. H. Jonassen & S. M. Land (Eds.), *Theoretical foundations of learning environments* (2nd ed., pp. 198–223). New York: Routledge. doi:10.4324/9780203813799

Blanton, M. L., & Stylianou, D. A. (2014). Understanding the role of transactive reasoning in classroom discourse as students learn to construct proofs. *The Journal of Mathematical Behavior*, *34*, 76–98. doi:10.1016/j.jmathb.2014.02.001

Bodemer, D., Ploetzner, R., Bruchmüller, K., & Häcker, S. (2005). Supporting learning with interactive multimedia through active integration of representations. *Instructional Science*, *33*(1), 73–95. doi:10.1007/s11251-004-7685-z

Borgers, N., de Leeuw, E., & Hox, J. (2000). Children as respondents in survey research: Cognitive development and response quality 1. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, *66*(1), 60–75. doi:10.1177/075910630006600106

Borgers, N., Hox, J., & Sikkel, D. (2003). Response quality in survey research with children and adolescents: The effect of labeled response options and vague quantifiers. *International Journal of Public Opinion Research*, *15*(1), 83–94. doi:10.1093/ijpor/15.1.83

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: Where young children go wrong. *Developmental Psychology*, *44*(5), 1478–1490. doi:10.1037/a0013110

Brooks, L. W., & Dansereau, D. F. (1983). Effects of structural schema training and text organization on expository prose processing. *Journal of Educational Psychology*, *75*(6), 811–820. doi:10.1037/0022-0663.75.6.811

Carr, J. M. (2012). Does math achievement h'APP'en when iPads and game-based learning are incorporated into fifth-grade mathematics instruction? *Journal of Information Technology Education: Research*, *11*, 269–286. doi:10.28945/1725

Chambers, C. T., & Johnston, C. (2002). Developmental differences in children's use of rating scales. *Journal of Pediatric Psychology*, *27*(1), 27–36. doi:10.1093/jpepsy/27.1.27

Chen, C. H., & Law, V. (2016). Scaffolding individual and collaborative game-based learning in learning performance and intrinsic motivation. *Computers in Human Behavior, 55*, 1201–1212. doi:10.1016/j.chb.2015.03.010

Cheng, R. W.-Y., Lam, S.-F., & Chan, J. C.-Y. (2008). When high achievers and low achievers work in the same group: The roles of group heterogeneity and processes in project-based learning. *British Journal of Educational Psychology*, *78*(2), 205–221. doi:10.1348/000709907X218160

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, *6*(3), 271–315. doi:10.1207/s15327809jls0603_1

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471–533. doi:10.1016/S0364-0213(01)00044-1

Chung, C. W., Lee, C. C., & Liu, C. C. (2013). Investigating face-to-face peer interaction patterns in a collaborative web discovery task: The benefits of a shared display. *Journal of Computer Assisted Learning*, *29*(2), 188–206. doi:10.1111/j.1365-2729.2012.00493.x

Clayphan, A., Collins, A., Kay, J., Slawitschka, N., & Horder, J. (2018). Comparing a single-touch whiteboard and a multi-touch tabletop for collaboration in school museum visits. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(1), 6:1-6:23. doi:10.1145/3191738

Cohen, E. G., & Lotan, R. A. (1995). Producing equal-status interaction in the heterogeneous classroom. *American Educational Research Journal, 32*(1), 99–120. doi:10.3102/00028312032001099

Coles, N. A., Larsen, J. T., & Lench, H. C. (2019). A meta-analysis of the facial feedback literature: Effects of facial feedback on emotional experience are small and variable. *Psychological Bulletin, 145*(6), 610-651. doi:10.1037/bul0000194

Colombo, L., & Landoni, M. (2014). A diary study of children's user experience with EBooks using flow theory as framework. *IDC '14: Proceedings of the 2014 Conference on Interaction Design and Children, Aarhus, Denmark*, 135–144. New York, NY: ACM. doi:10.1145/2593968.2593978

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*(2), 1047–1058. doi:10.1016/j.cognition.2007.04.010

Crook, C. (1998). Children as computer users: The case of collaborative learning. *Computers & Education*, *30*(3-4), 237–247. doi:10.1016/S0360-1315(97)00067-5

Curşeu, P. L., Schruijer, S. G. L., & Boroş, S. (2012). Socially rejected while cognitively successful? The impact of minority dissent on groups' cognitive complexity. *British Journal of Social Psychology*, *51*(4), 570–582. doi:10.1111/j.2044-8309.2011.02023.x

Curzon, P., & McOwan, P. W. (2018). *Computational Thinking: Die Welt des algorithmischen Denkens - in Spielen, Zaubertricks und Rätseln*. Berlin, Heidelberg, Germany: Springer. doi:10.1007/978-3-662-56774-6

Danish, J. A., Enyedy, N., Saleh, A., Lee, C., & Andrade, A. (2015). Science through technology enhanced play: Designing to support reflection through play and embodiment. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 1* (pp. 332–339). Gothenburg, Sweden: The International Society of the Learning Sciences. Abstract retrieved from https://www.isls.org/cscl2015/

Das digitale Klassenzimmer. (2018, December 06). *Zeit Online*. Retrieved from https://www.zeit.de/news/2018-12/06/das-digitale-klassenzimmer-181206-99-113958

Davidsen, J., & Ryberg, T. (2017). "This is the size of one meter": Children's bodily-material collaboration. *International Journal of Computer-Supported Collaborative Learning*, *12*(1), 65–90. doi:10.1007/s11412-017-9248-8

Deater-Deckard, K., Chang, M., & Evans, M. E. (2013). Engagement states and learning from educational games. *Digital Games: A Context for Cognitive Development. New Directions for Child and Adolescent Development, 139*, 21–30. doi:10.1002/cad.20028

Deater-Deckard, K., Chen, N., Wang, Z., McClelland, E., Nguyen, C., & Yamine, J. (2011). *Engagement States Observational Coding System (eSOCS)*. Blacksburg, VA: Virginia Tech. Retrieved from http://ltrg.centers.vt.edu/projects/games/Docs/eSOCS.pdf

Deater-Deckard, K., El Mallah, S., Chang, M., Evans, M. A., & Norton, A. (2014). Student behavioral engagement during mathematics educational video game instruction with 11-14 year olds. *International Journal of Child-Computer Interaction*, *2*(3), 101–108. doi:10.1016/j.ijcci.2014.08.001

de la Torre, J., Tjoe, H., Rhoads, K., & Lam, D. (2013). Conceptual and theoretical issues in proportional reasoning. *International Journal for Studies in Mathematics Education*, *6*(1), 21–38.

De Leeuw, E. D., & Otter, M. E. (1995). The reliability of children's responses to questionnaire items; question effects in children's questionnaire data. In J. J. C. M. Hox, B. F. van der Meulen, J. M. A. M. Janssens, L. W. C. Tavecchio, & J. J. F. ter Laak (Eds.), *Advances in family research* (pp. 251–257). Amsterdam, Netherlands: Thesis Publishers. Retrieved from https://hdl.handle.net/11245/1.116564

Demetriadis, S. N., Papadopoulos, P. M., Stamelos, I. G., & Fischer, F. (2008). The effect of scaffolding students' context-generating cognitive activity in technology-enhanced case-based learning. *Computers & Education*, *51*(2), 939–954. doi:10.1016/j.compedu.2007.09.012

Deutsche Telekom Stiftung. (2015). *Schule digital - Der Länderindikator 2015*. Retrieved from https://www.telekom-stiftung.de/sites/default/files/files/media/publications/schuledigital_2015_web.pdf

Die Bundesregierung. (2019). *Bundesrat stimmt für Grundgesetzänderung - Grünes Licht für "Digitalpakt Schule"*. Retrieved from https://www.bundesregierung.de/breg-de/themen/digitalisierung/gruenes-licht-fuer-digitalpakt-schule--1582804

Dillenbourg, P. (1999). What do you mean by collaborative learning? In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 1–19). Oxford, UK: Elsevier. Retrieved from https://telearn.archives-ouvertes.fr/hal-00190240

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL?* (pp. 61–91). Heerlen, Netherlands: Open Universiteit Nederland. Retrieved from https://telearn.archives-ouvertes.fr/hal-00190230

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning: From design to orchestration. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-enhanced learning: Principles and products* (pp. 3–19). Dordrecht, Netherlands: Springer Netherlands. doi:10.1007/978-1-4020-9827-7_1

Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, *32*(1), 268–291. doi:10.3102/0091732X07309371

Edelmann, S. (2008). *Computing the mind: How the mind really works*. Oxford, UK: Oxford University Press.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3-4), 169–200. doi:10.1080/02699939208411068

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

Ertl, B., Fischer, F., & Mandl, H. (2006). Conceptual and socio-cognitive support for collaborative learning in videoconferencing environments. *Computers & Education*, *47*(3), 298–315. doi:10.1016/j.compedu.2004.11.001

Falloon, G., & Khoo, E. (2014). Exploring young students' talk in iPad-supported collaborative learning environments. *Computers & Education*, *77*, 13–28. doi:10.1016/j.compedu.2014.04.008

Fleck, R., Rogers, Y., Yuill, N., Marshall, P., Carr, A., Rick, J., & Bonnett, V. (2009). Actions speak loudly with words: Unpacking collaboration around the table. *ITS '09: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, Banff, Canada,* 189–196. New York, NY: ACM. doi:10.1145/1731903.1731939

Franco, G. M., Muis, K. R., Kendeou, P., Ranellucci, J., Sampasivam, L., & Wang, X. (2012). Examining the influences of epistemic beliefs and knowledge representations on cognitive processing and conceptual change when learning physics. *Learning and Instruction*, *22*(1), 62–77. doi:10.1016/j.learninstruc.2011.06.003

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*(5), 349–358. doi:10.1037/0003-066X.43.5.349

Fuchs, M. (2005). Children and adolescents as respondents. Experiments on question order, response order, scale effects and the effect of numeric values associated with response options. *Journal of Official Statistics*, *21*(4), 701–725.

Fujimura, N. (2001). Facilitating children's proportional reasoning: A model of reasoning processes and effects of intervention on strategy change. *Journal of Educational Psychology*, *93*(3), 589–603. doi:10.1037/0022-0663.93.3.589

Furberg, A. (2016). Teacher support in computer-supported lab work: Bridging the gap between lab experiments and students' conceptual understanding. *International Journal of Computer-Supported Collaborative Learning*, *11*(1), 89–113. doi:10.1007/s11412-016-9229-3

Gallese, V., & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology*, *22*(3-4), 455–479. doi:10.1080/02643290442000310

Gelmini-Hornsby, G., Ainsworth, S., & O'Malley, C. (2011). Guided reciprocal questioning to support children's collaborative storytelling. *International Journal of Computer-Supported Collaborative Learning*, *6*(4), 577–600. doi:10.1007/s11412-011-9129-5

Gijlers, H., Weinberger, A., van Dijk, A. M., Bollen, L., & van Joolingen, W. (2013). Collaborative drawing on a shared digital canvas in elementary science education: The effects of script and task awareness support. *International Journal of Computer-Supported Collaborative Learning, 8*(4), 427–453. doi:10.1007/s11412-013-9180-5

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 586–596. doi:10.1002/wcs.55

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*(6), 516–522. doi:10.1111/1467-9280.00395

Gomoll, A. S., Hmelo-Silver, C. E., Tolar, E., Šabanović, S., & Francisco, M. (2017). Moving apart and coming together: Discourse, engagement, and deep learning. *Jounal of Educational Technology & Society*, *20*(4), 219–232. Retrieved from https://www.jstor.org/stable/26229219

Greene, B. A. (2015). Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist*, *50*(1), 14–30. doi:10.1080/00461520.2014.989230

Gweon, G., Jain, M., McDonough, J., Raj, B., & Rosé, C. P. (2013). Measuring prevalence of other-oriented transactive contributions using an automated measure of speech style accommodation. *International Journal of Computer-Supported Collaborative Learning*, *8*(2), 245–265. doi:10.1007/s11412-013-9172-5

Haas, C. (2018, June 03). Zwischen Mobbing und Medienkompetenz. *Zeit Online*. Retrieved from https://www.zeit.de

Hall, L., Hume, C., & Tazzyman, S. (2016). Five Degrees of Happiness: Effective smiley face Likert scales for evaluating with children. *IDC '16: Proceedings of the 15th International Conference on Interaction Design and Children, Manchester, UK*, 311–321. New York, NY: ACM. doi:10.1145/2930674.2930719

Hamilton, B. H., Nickerson, J. A., & Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy*, *111*(3), 465–497. doi:10.1086/374182

Harley, J. M., Bouchet, F., Hussain, M. S., Azevedo, R., & Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior*, *48*, 615–625. doi:10.1016/j.chb.2015.02.013

Hassan, M. M., Bashir, S., & Mussel, P. (2015). Personality, learning, and the mediating role of epistemic curiosity: A case of continuing education in medical physicians. *Learning and Individual Differences*, *42*, 83–89. doi:10.1016/j.lindif.2015.07.018

Hayes, A. F., & Rockwood, N. J. (2017). Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour Research and Therapy*, *98*, 39–57. doi:10.1016/j.brat.2016.11.001

Himmer, N. (2015, February 10). Das wischende Klassenzimmer. *Frankfurter Allgemeine*. Retrieved from https://www.faz.net/aktuell/beruf-chance/campus/unterricht-mit-dem-tablet-13412364.html

Hiniker, A., Sobel, K., Hong, S. R., Suh, H., Irish, I., Kim, D., & Kientz, J. A. (2015). Touchscreen prompts for preschoolers: Designing developmentally appropriate techniques for teaching young children to perform gestures. *IDC '15: Proceedings of the 14th International Conference on Interaction Design and Children, Boston, MA,* 109–118. New York, NY: ACM. doi:10.1145/2771839.2771851

Hulse, T., Daigle, M., Manzo, D., Braith, L., Harrison, A., & Ottmar, E. (2019). From here to there! Elementary: A game-based approach to developing number sense and early algebraic understanding. *Educational Technology Research and Development*, *67*(2), 423–441. doi:10.1007/s11423-019-09653-8

iPad 2 - Technical Specifications. (2013, June 14). Retrieved from https://support.apple.com/kb/SP622?viewlocale=en_US

Isaacs, B. (2018). *Understanding the Montessori approach: Early years education in practice* (2nd ed.). New York: Routledge. doi:10.4324/9781315536880

Isohätälä, J., Näykki, P., Järvelä, S., & Baker, M. J. (2018). Striking a balance: Socio-emotional processes during argumentation in collaborative learning interaction. *Learning, Culture and Social Interaction*, *16*, 1–19. doi:10.1016/j.lcsi.2017.09.003

Jackson, S. L., Krajcik, J., & Soloway, E. (1998). The design of guided learner-adaptable scaffolding in interactive learning environments. *CHI '98: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Los Angeles, CA*, 187–194. New York, NY: ACM Press/Addison-Wesley Publishing Co. doi:10.1145/274644.274672

Jakobsen, M. R., & Hornbæk, K. (2016). Negotiating for space? Collaborative work using a wall display with mouse and touch input. *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA*, 2050–2061. New York, NY: ACM. doi:10.1145/2858036.2858158

Jamil, I., Montero, C. S., Perry, M., O'Hara, K., Karnik, A., Pihlainen, K., … Subramanian, S. (2017). Collaborating around digital tabletops: Children's physical strategies from India, the UK and Finland. *ACM Transactions on Computer-Human Interaction*, *24*(3), 23:1-23:30. doi:10.1145/3058551

Janssen, J., Erkens, G., & Kirschner, P. A. (2011). Group awareness tools: It's what you do with it that matters. *Computers in Human Behavior*, *27*(3), 1046–1058. doi:10.1016/j.chb.2010.06.002

Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist*, *48*(1), 25–39. doi:10.1080/00461520.2012.748006

Järvenoja, H., & Järvelä, S. (2005). How students describe the sources of their emotional and motivational experiences during the learning process: A qualitative approach. *Learning and Instruction*, *15*(5), 465–480. doi:10.1016/j.learninstruc.2005.07.012

Järvenoja, H., & Järvelä, S. (2009). Emotion control in collaborative learning situations: Do students regulate emotions evoked by social challenges? *The British Journal of Educational Psychology*, *79*(3), 463–481. doi:10.1348/000709909X402811

Jeong, H., & Hmelo-Silver, C. E. (2016). Seven affordances of computer-supported collaborative learning: How to support collaborative learning? How can technologies help? *Educational Psychologist*, *51*(2), 247–265. doi:10.1080/00461520.2016.1158654

Jitendra, A. K., Star, J. R., Rodriguez, M., Lindell, M., & Someki, F. (2011). Improving students' proportional thinking using schema-based instruction. *Learning and Instruction*, *21*(6), 731–745. doi:10.1016/j.learninstruc.2011.04.002

Johnson, R., Shum, V., Rogers, Y., & Marquardt, N. (2016). Make or shake: An empirical study of the value of making in learning about computing technology. *IDC '16: Proceedings of the 15th International Conference on Interaction Design and Children, Manchester, UK*, 440–451. New York, NY: ACM. doi:10.1145/2930674.2930691

Jorczak, R. L. (2011). An information processing perspective on divergence and convergence in collaborative learning. *International Journal of Computer-Supported Collaborative Learning*, *6*(2), 207–221. doi:10.1007/s11412-010-9104-6

Jordan, K. (2015). Massive open online course completion rates revisited: Assessment, length and attrition. *International Review of Research in Open and Distributed Learning*, *16*(3), 341–358. doi:10.19173/irrodl.v16i3.2112

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal*, *27*(4), 664–687. doi:10.3102/00028312027004664

Kirschner, F., Paas, F., & Kirschner, P. A. (2009). A cognitive load approach to collaborative learning: United brains for complex tasks. *Educational Psychology Review*, *21*(1), 31–42. doi:10.1007/s10648-008-9095-2

Knörzer, L., Brünken, R., & Park, B. (2016). Emotions and multimedia learning: The moderating role of learner characteristics. *Journal of Computer Assisted Learning*, *32*(6), 618–631. doi:10.1111/jcal.12158

Krippendorff, K. (2012). *Content Analysis: An Introduction to Its Methodology* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Kuhn, D., & Crowell, A. (2011). Dialogic argumentation as a vehicle for developing young adolescents' thinking. *Psychological Science*, *22*(4), 545–552. doi:10.1177/0956797611402512

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago, IL: The University of Chicago Press.

Lehman, B., Matthews, M., D'Mello, S., & Person, N. (2008). What are you feeling? Investigating student affective states during expert human tutoring sessions. In B. P. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent Tutoring Systems* (pp. 50–59). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-69132-7_10

Ligorio, M. B., & Van der Meijden, H. (2008). Teacher guidelines for cross-national virtual communities in primary education. *Journal of Computer Assisted Learning*, *24*(1), 11–25. doi:10.1111/j.1365-2729.2007.00240.x

Liu, C.-C., Chung, C.-W., Chen, N.-S., & Liu, B.-J. (2009). Analysis of peer interaction in learning activities with personal handhelds and shared displays. *Journal of Educational Technology & Society*, *12*(3), 127–142. Retrieved from https://www.jstor.org/stable/jeductechsoci.12.3.127

Liu, C. C., & Tsai, C. C. (2006). An analysis of peer interaction patterns as discoursed by on-line small group problem-solving activity. *Computers & Education, 50*(3), 627-639. doi:10.1016/j.compedu.2006.07.002

Looije, R., Neerincx, M. A., & de Lange, V. (2008). Children's responses and opinion on three bots that motivate, educate and play. *Journal of Physical Agents*, *2*(2), 13–20. doi:10.14198/JoPha.2008.2.2.03

Lund, K., Niccolai, G., Lavoué, E., Hmelo-Silver, C., Gweon, G., & Baker, M. (Eds.). (2019). *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported*

Collaborative Learning (CSCL) 2019, Volume 1. Lyon, France: International Society of the Learning Sciences. Retrieved from https://cscl2019.com/

Manske, S., Hecking, T., Chounta, I.-A., Werneburg, S., & Hoppe, H. U. (2015). Using differences to make a difference: A study on heterogeneity of learning groups. In O. Lindwall, P. Häkkinen, T. Koschmann, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 1* (pp. 182–189). Gothenburg, Sweden: The International Society of the Learning Sciences. Abstract retrieved from https://www.isls.org/cscl2015/

Marks, J., Copland, E., Loh, E., Sunstein, C. R., & Sharot, T. (2019). Epistemic spillovers: Learning others' political views reduces the ability to assess and use their expertise in nonpolitical domains. *Cognition*, 188, 74-84. doi:10.1016/j.cognition.2018.10.003

Marsh, H. W. (1986). Negative item bias in ratings scales for preadolescent children: A cognitive-developmental phenomenon. *Developmental Psychology*, *22*(1), 37–49. doi:10.1037/0012-1649.22.1.37

Marshall, P., Hornecker, E., Morris, R., Dalton, N. S., & Rogers, Y. (2008). When the fingers do the talking: A study of group participation with varying constraints to a tabletop interface. *3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems, Amsterdam, Netherlands*, 33–40. doi:10.1109/TABLETOP.2008.4660181

Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25–45. doi:10.1146/annurev.psych.57.102904.190143

Mason, L., Ariasi, N., & Boldrin, A. (2011). Epistemic beliefs in action: Spontaneous reflections about knowledge and knowing during online information searching and their influence on learning. *Learning and Instruction*, *21*(1), 137–151. doi:10.1016/j.learninstruc.2010.01.001

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and Emotion*, *23*(2), 209–237. doi:10.1080/02699930802204677

Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist*, *63*(8), 760–769. doi:10.1037/0003-066X.63.8.760

Mercer, N. (1994). The quality of talk in children's joint activity at the computer. *Journal of Computer Assisted Learning*, *10*(1), 24–32. doi:10.1111/j.1365-2729.1994.tb00279.x

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction*, *6*(4), 359–377. doi:10.1016/S0959-4752(96)00021-7

Mercer, N. (2004). Sociocultural discourse analysis: Analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, *1*(2), 137–168. doi:10.1558/japl.v1.i2.137

Mercier, E. M., & Higgins, S. E. (2013). Collaborative learning with multi-touch technology: Developing adaptive expertise. *Learning and Instruction*, *25*, 13–23. doi:10.1016/j.learninstruc.2012.10.004

Mercier, E. M., Higgins, S. E., & da Costa, L. (2014). Different leaders: Emergent organizational and intellectual leadership in children's collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning*, *9*(4), 397–432. doi:10.1007/s11412-014-9201-z

Mercier, H. (2016). The argumentative theory: Predictions and empirical evidence. *Trends in Cognitive Sciences*, *20*(9), 689–700. doi.:10.1016/j.tics.2016.07.001

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, … Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, *331*(6014), 176–182. doi:10.1126/science.1199644

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1999). Early fraction calculation ability. *Developmental Psychology*, *35*(1), 164–174. doi:10.1037/0012-1649.35.1.164

Moed, A., Otto, O., Pal, J., Singh, U. P., Kam, M., & Toyama, K. (2009). Reducing dominance in multiple-mouse learning activities. In C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *Computer Supported Collaborative Learning Practices - CSCL2009 Conference Proceedings* (pp. 360–364). Rhodes, Greece: International Society of the Learning Sciences. doi:10.3115/1600053.1600107

Moosbrugger, H., & Kelava, A. (2008). *Testtheorie und Fragebogenkonstruktion*. Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-71635-8

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*(2), 358–368. doi:10.1037/0022-0663.91.2.358

Muis, K. R., Pekrun, R., Sinatra, G. M., Azevedo, R., Trevors, G., Meier, E., & Heddy, B. C. (2015). The curious case of climate change: Testing a theoretical model of epistemic beliefs, epistemic emotions, and complex learning. *Learning and Instruction*, *39*, 168–183. doi:10.1016/j.learninstruc.2015.06.003

Nathan, M. J. (2012). Rethinking formalisms in formal education. *Educational Psychologist, 47*(2), 125–148. doi:10.1080/00461520.2012.667063

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9*(3), 184–211. doi:10.1207/s15327957pspr0903_1

Nielsen, J. (1993). *Usability engineering*. San Francisco, CA: Morgan Kaufmann Publishers Inc.

Noroozi, O., Teasley, S. D., Biemans, H. J. A., Weinberger, A., & Mulder, M. (2013). Facilitating learning in multidisciplinary groups with transactive CSCL scripts. *International Journal of Computer-Supported Collaborative Learning*, *8*(2), 189–223. doi:10.1007/s11412-012-9162-z

Ottmar, E., & Landy, D. (2017). Concreteness fading of algebraic instruction: Effects on learning. *Journal of the Learning Sciences, 26*(1), 51–78. doi:10.1080/10508406.2016.1250212

Paulus, I., Gigout, F. J., Kaczmarek, C., & Wydra, G. (2011). *Bewegte Schule - Leitfaden zur Bewegungsförderung für Schulen im Saarland*. LandesArbeitsgemeinschaft für Gesundheitsförderung Saarland e.V. Retrieved from https://www.saarland.de/

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*(2), 91–105. doi:10.1207/S15326985EP3702_4

Phielix, C., Prins, F. J., & Kirschner, P. A. (2010). Awareness of group performance in a CSCL-environment: Effects of peer feedback and reflection. *Computers in Human Behavior*, *26*(2), 151–161. doi:10.1016/j.chb.2009.10.011

Polo, C., Lund, K., Plantin, C., & Niccolai, G. P. (2016). Group emotions: The social and cognitive functions of emotions in argumentation. *International Journal of Computer-Supported Collaborative Learning*, *11*(2), 123–156. doi:10.1007/s11412-016-9232-8

Pontecorvo, C., & Girardet, H. (1993). Arguing and reasoning in understanding historical topics. *Cognition and Instruction*, *11*(3-4), 365–395. doi:10.1080/07370008.1993.9649030

Raes, A., Schellens, T., De Wever, B., & Benoit, D. F. (2016). Promoting metacognitive regulation through collaborative problem solving on the web: When scripting does not work. *Computers in Human Behavior*, *58*, 325–342. doi:10.1016/j.chb.2015.12.064

Read, J. C. (2008). Validating the Fun Toolkit: An instrument for measuring children's opinions of technology. *Cognition, Technology & Work*, *10*(2), 119–128. doi:10.1007/s10111-007-0069-9

Read, J. C., & MacFarlane, S. (2006). Using the fun toolkit and other survey methods to gather opinions in child computer interaction. *IDC '06: Proceedings of the 2006 Conference on Interaction Design and Children, Tampere, Finland*, 81–88. New York, NY: ACM. doi:10.1145/1139073.1139096

Regionalverband Saarbrücken. (2012). *1. Bildungsbericht für den Regionalverband Saarbrücken*. Saarbrücken, Germany: Regionalverband Saarbrücken.

Reinholz, D. L., Trninic, D., Howison, M., & Abrahamson, D. (2010). It's not easy being green: Embodied artifacts and the guided emergence of mathematical meaning. In P. Brosnan, D. B. Erchick, & L. Flevares (Eds.), *Proceedings of the 32nd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education: Optimizing Student Understanding in Mathematics* (pp. 1488–1496). Columbus, OH: The Ohio State University. Retrieved from https://www.pmena.org/

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, *13*(3), 273–304. doi:10.1207/s15327809jls1303_2

Rick, J. (2012a). Proportion: A tablet app for collaborative learning. *IDC '12: Proceedings of the 11th International Conference on Interaction Design and Children, Bremen, Germany,* 316–319. New York, NY: ACM. doi:10.1145/2307096.2307155

Rick, J. (2012b, July). *Tiny tabletops: A research agenda*. Position paper presented at a workshop: Digital Ecosystems for Collaborative Learning (DECL 2012): Embedding Personal and Collaborative Devices to Support Classrooms of the Future, Sydney, Australia. Abstract retrieved from https://sites.google.com/site/iclsdecl/papers

Rick, J., Bejan, A., Roche, C., & Weinberger, A. (2012). Proportion: Learning proportional reasoning together. In A. Ravenscroft, S. Lindstaedt, C. D. Kloos, & D. Hernández-Leo (Eds.), *21st Century Learning for 21st Century Skills: EC-TEL 2012* (pp. 513–518). Berlin, Heidelberg, Germany: Springer. doi:10.1007/978-3-642-33263-0_52

Rick, J., Kopp, D., Schmitt, L., & Weinberger, A. (2015). Tarzan and Jane share an iPad. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 1* (pp. 356–363). Gothenburg, Sweden: The International Society of the Learning Sciences. Abstract retrieved from https://www.isls.org/cscl2015/

Rick, J., Rogers, Y., Haig, C., & Yuill, N. (2009). Learning by doing with shareable interfaces. *Children, Youth and Environments*, *19*(1), 320–341.

Roschelle, J., Rafanan, K., Bhanot, R., Estrella, G., Penuel, B., Nussbaum, M., & Claro, S. (2010). Scaffolding group explanation and feedback with handheld technology: Impact on students' mathematics learning. *Educational Technology Research and Development*, *58*(4), 399–419. doi:10.1007/978-3-642-85098-1_5

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer Supported Collaborative Learning* (pp. 69–97). Berlin, Heidelberg, Germany: Springer. doi:10.1145/130893.952914

Rosebery, A. S., Ogonowski, M., DiSchino, M., & Warren, B. (2010). "The coat traps all your body heat": Heterogeneity as fundamental to learning. *Journal of the Learning Sciences*, *19*(3), 322–357. doi:10.1080/10508406.2010.491752

Saarland Ministerium für Bildung, Familie, Frauen und Kultur. (2009). *Kernlehrplan Mathematik Grundschule*. Retrieved from https://www.saarland.de/dokumente/thema_bildung/KLPGSMathematik.pdf

Sakr, M., Jewitt, C., & Price, S. (2014). The semiotic work of the hands in scientific enquiry. *Classroom Discourse*, *5*(1), 51–70. doi:10.1080/19463014.2013.868078

Sakr, M., Jewitt, C., & Price, S. (2016). Mobile experiences of historical place: A multimodal analysis of emotional engagement. *Journal of the Learning Sciences*, *25*(1), 51–92. doi:10.1080/10508406.2015.1115761

Sapolsky, R. M. (2018). Doubled-edged swords in the biology of conflict. *Frontiers in Psychology*, *9*, 1–9. doi:10.3389/fpsyg.2018.02625

Sapounidis, T., Demetriadis, S., Papadopoulos, P. M., & Stamovlasis, D. (2019). Tangible and graphical programming with experienced children: A mixed methods analysis. *International Journal of Child-Computer Interaction*, *19*, 67–78. doi:10.1016/j.ijcci.2018.12.001

Schellens, T., Van Keer, H., De Wever, B., & Valcke, M. (2007). Scripting by assigning roles: Does it improve knowledge construction in asynchronous discussion groups? *International Journal of Computer-Supported Collaborative Learning*, *2*(2-3), 225–246. doi:10.1007/s11412-007-9016-2

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning*, *5*(1), 43–102. doi:10.1007/s11412-009-9080-x

Schmitt, L. J. (2015). *Individual vs. collaborative learning with an iPad application to foster proportional understanding* (Unpublished master's thesis). Saarland University, Saarbrücken, Germany.

Schmitt, L. J., Rick, J., & Weinberger, A. (2019). BiCo: A bipolar continuous rating scale for children's technology evaluation. *Technology, Pedagogy and Education, 28*(5), 503–516. doi:10.1080/1475939X.2019.1661279

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Schmitt, L. J., & Weinberger, A. (2018). Computer-supported collaborative learning: Mediated and co-present forms of learning together. In J. Voogt, G. Knezek, R. Christensen, & K.-W. Lai (Eds.), *Second handbook of information technology in primary and secondary education* (pp. 217-231). Cham, Switzerland: Springer. doi:10.1007/978-3-319-71054-9_15

Schmitt, L. J., & Weinberger, A. (2019). Fourth graders' dyadic learning on multi-touch interfaces - versatile effects of verbalization prompts. *Educational Technology Research and Development, 67*(3), 519-539. doi:10.1007/s11423-018-9619-5

Schneps, M. H., Ruel, J., Sonnert, G., Dussault, M., Griffin, M., & Sadler, P. M. (2014). Conceptualizing astronomical scale: Virtual simulations on handheld tablet computers reverse misconceptions. *Computers & Education*, *70*, 269–280. doi:10.1016/j.compedu.2013.09.001

Schoenfeld, A. H. (1987). What's all the fuss about metacognition? In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 189–215). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Schooler, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Applied Cognitive Psychology*, *16*(8), 989–997. doi:10.1002/acp.930

Schoor, C., Narciss, S., & Körndle, H. (2015). Regulation during cooperative and collaborative learning: A theory-based review of terms and concepts. *Educational Psychologist*, *50*(2), 97–119. doi:10.1080/00461520.2015.1038540

Schukajlow, S., Kolter, J., & Blum, W. (2015). Scaffolding mathematical modelling with a solution plan. *ZDM Mathematics Education*, *47*(7), 1241–1254. doi:10.1007/s11858-015-0707-2

Scott, J. (2000). Children as respondents: The challenge for quantitative methods. In P. Christensen & A. James (Eds.), *Research with Children: Perspectives and Practices* (pp. 98–119). London, UK, New York, NY: RoutledgeFalmer.

Sharples, M., Scanlon, E., Ainsworth, S., Anastopoulou, S., Collins, T., Crook, C., … O'Malley, C. (2015). Personal inquiry: Orchestrating science investigations within and beyond the classroom. *Journal of the Learning Sciences*, *24*(2), 308–341. doi:10.1080/10508406.2014.944642

Sim, G., & Horton, M. (2012). Investigating children's opinions of games: Fun Toolkit vs. This or That. *IDC '12: Proceedings of the 11th International Conference on Interaction Design and Children, Bremen, Germany*, 70–77. New York, NY: ACM. doi:10.1145/2307096.2307105

Sim, G., MacFarlane, S., & Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. *Computers & Education*, *46*(3), 235–248. doi:10.1016/j.compedu.2005.11.021

Soussignan, R. (2002). Duchenne smile, emotional experience, and autonomic reactivity: A test of the facial feedback hypothesis. *Emotion*, *2*(1), 52–74. doi:10.1037//1528-3542.2.1.52

Spackman, J. S., & Yanchar, S. C. (2014). Embodied cognition, representationalism, and mechanism: A review and analysis. *Journal for the Theory of Social Behaviour*, *44*(1), 46–79. doi:10.1111/jtsb.12028

Stahl, G. (2005). Group cognition in computer assisted collaborative learning. *Journal of Computer Assisted Learning*, *21*(2), 79–90. doi:10.1111/j.1365-2729.2005.00115.x

Stahl, G. (2006). Supporting group cognition in an online math community: A cognitive tool for small-group referencing in text chat. *Journal of Educational Computing Research*, *35*(2), 103–122. doi:10.2190/Q435-7611-2561-720P

Stahl, G. (2013). Transactive discourse in CSCL. *International Journal of Computer-Supported Collaborative Learning*, *8*(2), 145–147. doi:10.1007/s11412-013-9171-6

Stahl, G. (2016). From intersubjectivity to group cognition. *Computer Supported Cooperative Work*, *25*(4-5), 355–384. doi:10.1007/s10606-016-9243-z

Statistisches Bundesamt. (2018). *Wirtschaftsrechnungen: Einkommens- und Verbrauchsstichprobe, Ausstattung privater Haushalte mit ausgewählten Gebrauchsgütern und Versicherungen* (Fachserie 15 Heft 1). Statistisches Bundesamt (Destatis). Retrieved from https://www.destatis.de/

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

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology*, *54*(5), 768–777. doi:10.1037/0022-3514.54.5.768

Tabor, B. K. (2000). At Zoo, orangutans use symbols, labels to communicate with humans. *Smithsonian Institution Research Reports, 102,* 1–6.

Teasley, S. D. (1997). Talking about reasoning: How important is the peer in peer collaboration? In L. B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse,* *tools and reasoning: Essays on situated cognition* (pp. 361–384). Berlin, Heidelberg, Germany: Springer. doi:10.1007/978-3-662-03362-3_16

ter Vrugte, J., de Jong, T., Vandercruysse, S., Wouters, P., van Oostendorp, H., & Elen, J. (2015). How competition and heterogeneous collaboration interact in prevocational game-based mathematics education. *Computers & Education*, *89*, 42–52. doi:10.1016/j.compedu.2015.08.010

Thornborrow, J. (2003). The organization of primary school children's on-task and off-task talk in a small group setting. *Research on Language and Social Interaction*, *36*(1), 7–32. doi:10.1207/S15327973RLSI3601_2

Tourniaire, F., & Pulos, S. (1985). Proportional reasoning: A review of the literature. *Educational Studies in Mathematics*, *16*(2), 181–204. doi:10.1007/bf02400937

Trevors, G. J., Muis, K. R., Pekrun, R., Sinatra, G. M., & Muijselaar, M. M. L. (2017). Exploring the relations between epistemic beliefs, emotions, and learning from texts. *Contemporary Educational Psychology*, *48*, 116–132. doi:10.1016/j.cedpsych.2016.10.001

Tsai, C.-C., Ho, H. N. J., Liang, J.-C., & Lin, H.-M. (2011). Scientific epistemic beliefs, conceptions of learning science and self-efficacy of learning science among high school students. *Learning and Instruction*, *21*(6), 757–769. doi:10.1016/j.learninstruc.2011.05.002

Tsovaltzi, D., Judele, R., Puhl, T., & Weinberger, A. (2015). Scripts, individual preparation and group awareness support in the service of learning in Facebook: How does CSCL compare to social networking sites? *Computers in Human Behavior*, *53*, 577–592. doi:10.1016/j.chb.2015.04.067

Tsovaltzi, D., Judele, R., Puhl, T., & Weinberger, A. (2017). Leveraging social networking sites for knowledge co-construction: Positive effects of argumentation structure, but premature knowledge consolidation after individual preparation. *Learning and Instruction*, *52*, 161–179. doi:10.1016/j.learninstruc.2017.06.004

Tsovaltzi, D., Weinberger, A., Schmitt, L., Bellhäuser, H., Müller, A., Konert, J., … van Aalst, J. (2019). Group Formation in the Digital Age: Relevant Characteristics, their Diagnosis, and Combination for Productive Collaboration. In K. Lund, G. Niccolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th*

International Conference on Computer Supported Collaborative Learning, Volume 2 (pp. 719-726). Lyon, France: The International Society of the Learning Sciences. Abstract retrieved from https://cscl2019.com/

van Boxtel, C., van der Linden, J., & Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. *Learning and Instruction*, *10*(4), 311–330. doi:10.1016/S0959-4752(00)00002-5

Van Der Meijden, H., & Veenman, S. (2005). Face-to-face versus computer-mediated communication in a primary school setting. *Computers in Human Behavior*, *21*(5), 831–859. doi:10.1016/j.chb.2003.10.005

van Dijk, A. M., Gijlers, H., & Weinberger, A. (2014). Scripted collaborative drawing in elementary science education. *Instructional Science, 42*(3), 353–372. doi:10.1007/s11251-013-9286-1

Vogel, F., Reichersdorfer, E., Kollar, I., Ufer, S., Reiss, K., & Fischer, F. (2013). Learning to argue in mathematics: Effects of heuristic worked examples and CSCL scripts on transactive argumentation. In N. Rummel, M. Kapur, M. Nathan, & S. Puntambekar (Eds.), *To See the World and a Grain of Sand: Learning Across Levels of Space, Time and Scale: CSCL 2013 Conference Proceedings Volume 1 – Full Papers & Symposia* (pp. 526–533). Madison, WI: The International Society of the Learning Sciences. Abstract retrieved from https://www.isls.org/cscl/2013/

Vogel, F., & Weinberger, A. (2018). Quantifying qualities of collaborative learning processes. In F. Fischer, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (Eds.), *International handbook of the learning sciences* (pp. 500–510). New York, NY: Routledge. doi:10.4324/9781315617572

Vygotsky, L. S. (1962). *Studies in communication: Thought and language* (E. Hanfmann & Gertrude Vakar, Eds.). Cambridge, MA: MIT Press. doi:10.1037/11193-000

Webb, N. M. (1982). Group composition, group interaction, and achievement in cooperative small groups. *Journal of Educational Psychology*, *74*(4), 475–484. doi:10.1037/0022-0663.74.4.475

Webb, N. M. (1989). Peer interaction and learning in small groups. *International Journal of Educational Research*, *13*(1), 21–39. doi:10.1016/0883-0355(89)90014-1

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

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

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

Weinberger, A., Hartmann, C., Schmitt, L. J., & Rummel, N. (2018). Computer-unterstützte kooperative Lernszenarien. In H. Niegemann & A. Weinberger (Eds.), *Lernen mit Bildungstechnologien: Praxisorientiertes Handbuch zum intelligenten Umgang mit digitalen Medien* (pp. 1-19*)*. Berlin, Heidelberg: Springer. doi:10.1007/978-3-662-54373-3_20-1

Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning and Instruction*, *17*(4), 416–426. doi:10.1016/j.learninstruc.2007.03.007

Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior*, *26*(4), 506–515. doi:10.1016/j.chb.2009.08.007

Weinberger, A., Stegmann, K., Fischer, F., & Mandl, H. (2007). Scripting argumentative knowledge construction in computer-supported learning environments. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting computer-supported collaborative learning: Cognitive, computational and educational perspectives* (pp. 191–211). Boston, MA: Springer. doi:10.1007/978-0-387-36949-5_12

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–636. doi:10.3758/BF03196322

Wosnitza, M., & Volet, S. (2012). Group heterogeneity and homogeneity in personal content goals for a group learning activity: Impact on individual appraisals. *Applied Psychology*, *61*(4), 585–604. doi:10.1111/j.1464-0597.2012.00507.x

Wouters, P., van Oostendorp, H., ter Vrugte, J., Vandercruysse, S., de Jong, T., & Elen, J. (2015). The role of curiosity-triggering events in game-based learning for mathematics.

In J. Torbeyns, E. Lehtinen, & J. Elen (Eds.), *Describing and studying domain-specific serious games* (pp. 191–207). Cham, Switzerland: Springer. doi:10.1007/978-3-319-20276-1

Zaman, B., Abeele, V. V., & De Grooff, D. (2013). Measuring product liking in preschool children: An evaluation of the Smileyometer and This or That methods. *International Journal of Child-Computer Interaction*, *1*(2), 61–70. doi:10.1016/j.ijcci.2012.12.001

Zuckerman, O., Arida, S., & Resnick, M. (2005). Extending tangible interfaces for education: Digital Montessori-inspired manipulatives. *CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Portland, OR*, 859-868. New York, NY: ACM. doi:10.1145/1054972.1055093

Appendix: Publications

Study I:

Schmitt, L. J., & Weinberger, A. (2018). Computer-supported collaborative learning: Mediated and co-present forms of learning together. In J. Voogt, G. Knezek, R. Christensen, & K.-W. Lai (Eds.), *Second handbook of information technology in primary and secondary education* (pp. 217–231). Cham, Switzerland: Springer. doi:10.1007/978-3-319-53803-7_15-1

Study II:

Schmitt, L. J., Rick, J., & Weinberger, A. (2019). BiCo: A bipolar continuous rating scale for children’s technology evaluation. *Technology, Pedagogy and Education, 28*(5), 503–516. doi:10.1080/1475939X.2019.1661279

Study III:

Schmitt, L. J., & Weinberger, A. (2019). Fourth graders' dyadic learning on multi-touch interfaces – versatile effects of verbalization prompts. *Educational Technology Research and Development, 67*(3), 519-539. doi:10.1007/s11423-018-9619-5

Study IV:

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Study I:

Schmitt, L. J., & Weinberger, A. (2018). Computer-supported collaborative learning: Mediated and co-present forms of learning together. In J. Voogt, G. Knezek, R. Christensen, & K.-W. Lai (Eds.), *Second handbook of information technology in primary and secondary education* (pp. 217–231). Cham, Switzerland: Springer. doi:10.1007/978-3-319-53803-7_15-1

Computer-Supported Collaborative Learning - Mediated and Co-Present Forms of Learning Together

Lara Johanna Schmitt and Armin Weinberger

Abstract

Computer-support offers versatile opportunities for learning together at a distance as well as for co-located scenarios of collaborative learning, in which learners construct arguments, share knowledge, and jointly produce task solutions themselves. In this chapter, we explore different computer-mediated and co-present forms of computer-supported collaborative learning (CSCL). We will discuss both the potential as well as the problems of CSCL. Finally, we will show how current instructional approaches aim to overcome CSCL problems to realize CSCLs full potential.

Keywords: CSCL, Computer-Supported Collaborative Learning, co-present, computer-mediated

A straightforward and concise way of defining Computer-Supported Collaborative Learning (CSCL) is to state that it involves some form of learning together with peers and computer-based support for learning. Starting from there, however, multiple approaches to collaborative learning can be distinguished, and computer support can vary considerably and is rapidly progressing along with information and communication technologies (ICT).

Collaborative learning. One of the most widely disseminated and intuitive distinctions of learning together with peers is the one between collaborative and cooperative learning. A well-established definition of collaboration is "Collaboration is a coordinated, synchronous activity that is the result of continued attempt to construct and maintain a shared conception of a problem" (Roschelle & Teasley, 1995, p. 70); thus, collaborative learning is different from cooperative learning, in which a group of learners splits the work into individual, additive sub-tasks. While per definition being a "synchronous activity" (Roschelle & Teasley, 1995, p. 70), asynchronous CSCL can still host collaboration with learners' reasoning processes being not temporally but transactively intertwined (Roschelle & Teasley, 1995), i.e. learners address and build on the reasoning of their peers. Respectively, the "interactions among peers" (Dillenbourg, Järvelä, & Fischer, 2009, p. 3) were termed as the main feature of collaborative learning. Furthermore, learners can convergence in collaboration and share knowledge as a result (Weinberger, Stegmann, & Fischer, 2007). One of the foci in investigating collaborative learning is to analyze how learners acquire knowledge by engaging in specific argumentative discourse activities, which have been classified along different epistemic, argumentative and social dimensions (Weinberger & Fischer, 2006).

- Regarding the *epistemic quality* of discourse, learners may be more or less on-task, may demonstrate a more or less correct understanding of concepts, and may more or less adequately link and apply concepts when solving a problem case.
- *Argumentative quality* refers to how learners elaborate their line of reasoning in formal argumentative structures, including the extent learners warrant their claims, produce evidence, and anticipate and include counter-arguments in their line of reasoning.

— The *social modes of interaction* regard how learners refer to and build on each other's reasoning, which has been termed the degree of transactivity (Stahl, 2013). Here, learners could use peers as additional resource and elicit knowledge from each other, or agree and disagree with each other's stances.

Past research has shown that transactivity is a core quality of learners' discourse, strongly related to individual and group learning outcomes (e.g., Noroozi, Weinberger, Biemans, Mulder, & Chizari, 2013). A highly transactive act is, for instance, to relate to peer contributions in a critical way with the goal to sort out differences of perspectives by subsequent negotiation, integration, and modification of understanding. Transactivity also builds on the establishment of a common ground among collaborative learners (Beers, Boshuizen, Kirschner, & Gijselaers, 2006). Common ground refers to the externalization and transformation or re-contextualization of the group members' knowledge or "grounding", to reach mutual understanding (Bechky, 2003, p. 534), the base for all further interactions and knowledge construction processes (Bechky, 2003).

Specifically, in the context of early school education, the success of collaborative learning largely depends on a sense of intimacy and mutuality through sharing experiences and common goals (Crook, 1998). The concept of common ground relates to the question of how learners co-develop more or less similar knowledge structures when learning together and converge on knowledge. Knowledge convergence has been conceptualized and investigated as an important quality of collaborative learning (Weinberger et al., 2007). Through processes of establishing common ground and transactive discourse, learners may acquire equivalent amounts of knowledge and share knowledge as a result of learning together. This group level perspective on collaborative learning underlines the importance of epistemic, argumentative, and transactive qualities of discourse. While successful groups of learners spontaneously display elaborated and critical forms of consensus-building, there are robust findings that collaborative learners typically benefit from additional scaffolding promoting these discourse qualities and facilitating knowledge convergence (Kreijns, Kirschner, & Vermeulen, 2013; Noroozi et al., 2013). Hence, scaffolding learners to engage in productive interactions is a major focus in CSCL.

Computer-support for collaborative learning. Computers can be used to support remote learning as well as face-to-face learning settings (co-located CSCL), they enable asynchronous discussions, e.g., discussion boards, and synchronous learning and working, e.g. simultaneously drafting documents in shared workspaces together. Computers are tools for representing and sharing ideas and knowledge. Externalized knowledge that is being developed jointly, e.g. as a mind-map, can help learners to focus on important concepts and foster the quality of learner products as well as learning outcomes (Janssen, Erkens, Kirschner, & Kanselaar, 2010).

The following two sections introduce and discuss two main branches of CSCL, mediated and co-located computer-supported collaborative learning. We will give insights into exemplary learning environments and research done in this field, drawing on seminal work, constitutional to the research area of CSCL, as well as recent developments.

Computer-Mediated Collaborative Learning

Building on the 'communication'-function of ICT, computer-mediated learning entails all forms of learning together at a distance. Computer-mediated communication (CMC) can be relatively synchronous or asynchronous, that is participants in online interaction may use platforms, in which an immediate response is expected, such as in chats or video conferences, or may expect any kind of delay on platforms such as discussion boards or email. Moreover, communication may be more or less enriched by media formats building on platforms that support sharing of, for example, videos or text only.

Computer-mediated learning scenarios are often designed for online learning groups of adult learners, but can enable learning across schools (Ligorio & Van der Meijden, 2008), in blended scenarios of learning in and out of schools and combining online and face-to-face dialogue. Computer-mediated learning scenarios mainly address secondary school students and adults, which we will outline in the following paragraphs.

Typical scenarios of computer-mediated collaborative learning.

Computer media for communication span from asynchronous, text-based emails and discussion boards to synchronous, multi-media environments like video-conferencing and virtual worlds. Video is playing a major role in transmitting lectures or providing one-on-one tutorials, but only few CSCL studies have been focusing on video conferencing (e.g., Ertl, Fischer, & Mandl, 2006). In these scenarios, learners are communicating by video-chat and simultaneously work on a shared workspace, e.g., creating a concept map together or co-constructing arguments.

Asynchronous and text-based scenarios of CSCL, in contrast, have been widely investigated and applied in different platforms and contexts (Tsovaltzi, Judele, Puhl, & Weinberger, 2015). Learners in asynchronous environments have equal opportunities for participating in argumentative discourse by constructing arguments at their own time without being interrupted with the ability to search for and include additional resources. Small groups of online learners can be composed to jointly inquire scientific problems and phenomena, but also larger online courses and MOOCs (Massive Open Online Courses) do partly build on collaborative scenarios and typically provide a platform for social interaction. Likewise, CSCL in social networks is moving beyond small group interaction and is supporting collaboration among larger communities. Here, learners provide personal profiles and share information across different groups of friends and larger communities fostering social interaction (Kreijns et al., 2013). A recent study found how learning in social networks strongly builds on knowledge sharing and co-construction to the extent that individual preparation can lead to premature knowledge consolidation and impede learning gains in argumentative online learning scenarios (Tsovaltzi et al., 2015).

In other advanced, but under-investigated computer-mediated scenarios of CSCL, learners direct avatars through virtual landscapes, where they meet and interact with avatars of peers. These scenarios are strongly influenced by massive multi-player online games and often build on the respective game environments. For instance, *Second Life* is a game environment in which, for instance, foreign language learning is being investigated utilizing the virtual context as a visual and linguistic support (Hsiao, Yang, & Chu, 2015).

Platforms sometimes support different types of communication media. For instance, *Euroland* offers both asynchronous discussion boards and a virtual world, in which learners need to be carefully introduced to the tools and possible courses of action for expanding a local classroom into a larger online community (Ligorio & Van der Meijden, 2008). The seminal Virtual Math Teams project (Stahl, 2006) connects people over the Internet to collaboratively discuss and solve math problems using a shared whiteboard and a chat tool. This project shifted from a local intervention to building an online math community connecting students from different classrooms all over the world to participate in discourse for a deep understanding of mathematical principles (Stahl, 2006).

Issues of computer-mediated collaborative learning

Regardless of the communication media, there are indications that Computer-Mediated Communication (CMC) requires learners to more explicitly coordinate themselves than in face-to-face (FTF) environments (Van der Meijden & Veenman, 2005). The "poorer" the medium is, the less socially present peers are and the more pronounced are effects of the medium on different qualities of communication. So for instance, the effect of an increased need for coordination is more pronounced in text-based communication and mitigated, but still observable in "richer" video conferences (Ertl et al., 2006). Social context cues are what seem to be lost in poorer scenarios, such as discussion boards, in comparison to rich communication scenarios like FTF or video conferencing. At the dawn of CMC in the late last century, lack of social context cues has been related to problems of deindividuation, i.e. submergence in an anonymous group with a loss of identity and individual responsibility leading to violation of social norms online; however, later CMC models indicated how lack of social context cues may also increase group identity and foster social attraction (for an overview, see Walther, 2011).

Walther's hyperpersonal model of CMC exceeding face-to-face interaction in allowing to control, edit and optimize presentation of one's self (Walther, 1996, 2011) appears to be particularly relevant for communication in social networking sites (SNS) like Facebook, in which users can cultivate a profile and intentionally disclose their personality to a high degree. In this way, Facebook and other SNS may foster social aspects of academic life including actual friendships (e.g., Selwyn, 2009), which can contribute to study-related knowledge sharing (Wodzicki, Schwämmlein, & Moskaliuk, 2012). SNS can also greatly support active information seeking and building social capital that increases the connectedness with people for future support and cooperation (Lampe, Vitak, Gray, &

Ellison, 2012). Interestingly, media effects may not necessarily impede learning, because learners can compensate media effects over time or can be additionally supported to overcome media limitations. Hence, advantages of asynchronous CMC for learning - independent of place and time - also factor into how learners can harness CMC for higher levels of CSCL. In online discussion boards, learners have the time they individually need to inquire additional resources, construct elaborate arguments, and transactively build on the reasoning of their peers.

Another issue of computer-mediated collaborative learning is, however, how motivational problems, well-known in collaborative learning research, translate and potentially aggravate in computer-mediated scenarios. Online learning seems to particularly suffer from attrition and lack of participation. Up to half of learners drop out of regular online courses, a rate that can go up to 90% in MOOCs learners take by free choice (Jordan, 2015). Moreover, participation within online courses drops rapidly after the first week and often precedes drop-out (Nistor & Neubauer, 2010). Reasons for dropping-out are numerous and include personal motivational problems as well as technical difficulties (Park & Choi, 2009; Sitzmann, Ely, Bell, & Bauer, 2010). Registration to online courses can be easy enough for learners to have a peek at the learning material, but participation hard enough to quickly reconsider and refocus on alternative courses or learners do not intend to complete the course in the first place. Within dedicated CSCL scenarios that build on student agency and engagement, lack of and heterogeneity of participation have been discussed as particularly harmful to motivation and learning (Weinberger & Fischer, 2006). Motivationally detrimental effects in CSCL work on the group level, i.e. peers influence each other's motivation, and can be enhanced as learners may have difficulties to contact each other online, i.e. online learning partners can be more elusive than FTF learners. One well-known problem is free-riding, which refers to a team member not investing the required effort assuming that other team members will do (Kerr & Bruun, 1983) and vice versa the so-called ‘sucker-effect’ of one learner covering major parts of the task and hence, losing motivation. This suboptimal distribution of group work can tremendously reduce the potential of collaborative learning for equal participation in argumentative elaboration activities (Cohen & Lotan, 1995). While collaborative learning aims to establish equal participation and access to opportunities for argumentation and active learning, there are robust findings that collaborative learning often produces the exact opposite (e.g., Nihalani, Wilson, Thomas, & Robinson, 2010). Advanced learners benefit, in particular, from complex, collaborative learning scenarios, whereas participants with suboptimal preconditions for learning are left behind. Finally, it has been found that a sensible argumentation culture in online scenarios can not be expected as a short-term goal, but needs time, even weeks or months, to develop (Puhl, Tsovaltzi, & Weinberger, 2015).

Co-Present Computer-Supported Collaborative Learning

Contrary to computer-mediated collaborative learning discussed above, in co-present CSCL scenarios, learners are physically together, interacting with each other face-to-face and with one or several technologies in order to construct knowledge. Devices to support co-present collaborative learning are interactive whiteboards, desktop computers, hand-helds, tablets, or tabletops. The "social situation" positively affects learners' motivation and communication (Roschelle & Teasley, 1995, p. 69). In early co-present CSCL scenarios, e.g. Logo programming, learners typically had to share the access to the input devices (mouse, keyboard). Only one learner at a time could manipulate the representations on the screen, while the learning partner was cast back to a more passive role or the role of thinker. Roles and actions had to be negotiated. With the advent of multi-touch interfaces (tabletops, tablets), several learners of all age groups can simultaneously access and directly manipulate a shared representation. Thus, new ways of interacting with each other and with the learning environment emerged, which may be conducive to embodied cognition (Schneps et al., 2014). In the context of learning, embodiment refers to the role of the body in interacting with the (learning) environment to construct knowledge, and is gaining more and more attention in the learning sciences (Abrahamson, 2017).

Typical Scenarios of co-present collaborative learning

Primary education

Taking the role of embodiment into account, researchers have been designing for *collective* embodied learning experiences. For example, in the STEP (Science Through Technology Enhanced Play) project (Danish, Enyedy, Saleh, Lee, & Andrade, 2015), 7-8-year-old children playfully learn about chemical concepts by acting out particles. A projection on the wall depicts a chemical element that changes its matter of state depending on the children's (the particles') movements through the classroom.

CSCL among primary school children can also be realized around a multi-touch table, in which, for instance, students in groups of four can be encouraged to discuss complex tasks and organize and distribute responsibilities and roles within their groups (Mercier, Higgins, & da Costa, 2014). In another project utilizing iPads as mini-tabletops, fourth graders (on average 10 years old) collaborated in dyads on one shared tablet to work out proportional reasoning tasks with the Proportion app (Rick, Kopp, Schmitt, & Weinberger, 2015). Though tablets are personal devices, usually not used by several people at once, this scenario intentionally brings learners together. The pedagogical idea is to combine an embodied and immersive way of learning (moving two bars on the tablet to map volumes with numbers) with the benefits of collaboration (reasoning together). Peculiar interaction patterns can emerge in such a scenario: a case analysis of "Tarzan and Jane" (Rick et al., 2015), who were notably heterogeneous in their approach to learning, showed remarkable learning gains. Tarzan and Jane started with under-average pre-test results, but ended up being exceedingly successful in solving the Proportion tasks and acquiring knowledge. Beyond this specific case, significant learning gains were found in a sample of n=162 primary school children that collaboratively used Proportion (Schmitt & Weinberger, 2017), pointing at the benefits of embodied collaborative learning.

For a scenario to be classified as co-present collaborative learning, learners do not necessarily need to share one device. In the TechPALS (Technology-mediated, Peer-Assisted Learning) learning environment (Roschelle et al., 2010), fourth graders were working in groups of three solving math problems with everyone using their own hand-held device. However, the students had to work closely together, agree on solutions, and share resources in order to make progress, and feedback was only provided on the group level. The collaboratively working students gained significantly more domain knowledge compared to a control group working individually (Roschelle et al., 2010).

That said, collaborative and individual learning approaches are not mutually exclusive, but can be effectively orchestrated as parts of a larger learning unit. For example, in a study by Gijlers, Weinberger, van Dijk, Bollen, and van Joolingen (2013), children aged 10-11 years first created representations of a scientific phenomenon (photosynthesis) on separate tablets. Individually creating a first version of the drawing, was then followed up by sharing and comparing drawings with a peer and finding a joint solution. The orchestration of individual and collaborative learning phases showed to enhance the quality of dialogue and knowledge outcomes compared to a collaborative-only control group (Gijlers et al., 2013).

Secondary education

Shifting the focus to older students, the next paragraphs will focus on exemplary studies of co-present collaboration at the secondary education level.

Quite common are learning environments that enable simulating phenomena of natural sciences. In their seminal paper, Roschelle and Teasley (1995) delineate the learning processes of two 15-year-old collaborative learners who are acquiring an understanding of Newtonian physics with the “Envisioning Machine” that served as a joint problem space. “Co-Lab” is a scientific inquiry-learning environment that students use either individually or collaboratively (via internet using a chat-tool or face-to-face sharing a computer). It was found that providing high-school students’ (16-18 years) with a regulation tool (i.e. supporting planning, monitoring and evaluating) has positive effects on planning and evaluating activities, supposedly because of the complex and novel task the students had to carry out (Manlove, Lazonder, & de Jong, 2009).

The learning environment “WISE” (web-based inquiry science environment) was employed to help 16-year-old students in web-based inquiry learning (Raes, Schellens, De Wever, & Benoit, 2016). Groups of two shared one computer to search the Internet and collect information about the topics global warming and climate change. Central to the activity was critically assessing the quality of the information found on the Internet. To facilitate advantageous interactions, like equal participation between the students, a collaboration script was added.

Issues of co-present collaborative learning

Despite the advantage of interacting directly with each other, there are also issues of co-present collaborative learning. The challenges addressed in the following paragraphs add to the general requirements, like transactivity, regulation of learning, and task focus, that need to be met for effective collaborative learning.

One issue is the question of access to the environment and territoriality, i.e., the subdivision of available workspace in a jointly used device. The available space might either be open and accessible for everyone, so that learners would need to negotiate and co-regulate their input activities and times, or the instructions might steer towards specific interaction patterns (Dillenbourg & Evans, 2011). Designing for productive territorial interaction may be difficult for young learners. For example, in the Tarzan and Jane case study (Rick et al., 2015), peers could reach over and interfere with the partner's virtual objects. In spacious tabletop environments, each participant may have their own territory and input devices at their disposal. After having worked on a tabletop and produced input, e.g., when jointly constructing a mind-map, learners may, however, at times move altogether to one side of the tabletop, rotate all concepts to be read upright, look at and reference their map when discussing their work. Hence, usability issues are to be taken into account when designing co-present learning environments: How should learners share the workspace? To what extent does the workspace enable equal sharing of activities? Where should learners be positioned? Does everyone get their own input opportunity or working space?

Closely related to questions of territoriality are dominant interaction patterns in co-present CSCL. Especially in desktop computer settings, one learner might dominate the other(s) by controlling the mouse and not sharing access to the input device. Solutions proposed to address this issue are, for example, encouraging turn-taking and providing a split-screen (Moed et al., 2009), or providing several mice (Stanton, Neale, & Bayon, 2002), so that every learner would have the chance to interact with the learning environment, which may, however, render the learning scenario cooperative rather than collaborative. Providing several mice, albeit beneficial for student engagement and participation overall, fosters co-operative and parallel interactions (Stanton et al., 2002). Also, dominant behavior does not necessarily disappear when every learner has access to their own mouse (Moed et al., 2009). When multiple input opportunities via multi-touch vs. multiple mice on a tabletop were directly compared, multi-touch was found to support fluid interactions and reduce coordination effort compared to the multiple mice condition (Hornecker, Marshall, Dalton, & Rogers, 2008).

There are also concerns that playful, embodied learning environments might be overly hands-on, neglect verbalization and elaboration of concepts, and thus do not support abstraction to the underlying concepts. For example, in Danish et al.'s (2015) study where children acted out particles, the situation sometimes got a bit out of hand, and children would run around the classroom and even play tag. To prevent learners being "lost in action", sensibly combining hands-on environments with phases of reflection is needed (Danish et al., 2015), in order to improve the quality of learners' epistemic and social activities. However, reflection phases need to be introduced and intertwined carefully with the hands-on activities, to not interrupt learners' flow and disengage them (Schmitt & Weinberger, 2017).

Finally, the practical-technical and pedagogical issues of integrating computer-based tools into everyday classroom practices are a not to be underestimated challenge for teachers, calling for systematic professional development measures (Vanderlinde, Aesaert, & van Braak, 2014).

Solutions / Facilitating CSCL environments

The following paragraphs will discuss two fundamental strategies to support effective collaborative learning: structuring vs. regulating interactions (Dillenbourg, 2002).

Structuring interactions

One well-investigated method to structure social interactions in CSCL are collaboration scripts. These scripts assign learning activities and roles to the learners and reduce time spent on organizational issues and generally show a positive impact on learning (Scheuer, Loll, Pinkwart, & McLaren, 2010). Encouraging high-schoolers to develop joint group norms regarding their communication and task solving positively affected their collaborative interactions and learning outcomes (Zahn, Krauskopf, Hesse, & Pea, 2012). The Guided Reciprocal Peer Questioning script successfully fostered primary school children's collaborative discussions and quality of the learning product (Gelmini-Hornsby, Ainsworth, & O'Malley, 2011). The beneficial effects of scripting seem to be rather consistent, bearing in mind that too much scripting might lead to "over-scripting", meaning that an overly tight specification of interactions would ruin the nature of collaboration, negatively influencing interactions and acceptance of the script (Dillenbourg, 2002). Consequently, the question of amount, content, timing and general presentation of scripts is a complex one (see also, Fischer, Kollar, Stegmann, & Wecker, 2013). One of the future directions is seen in the development of adaptive scripts (Fischer et al., 2013).

Supporting learners' behavior in computer-supported learning environments is needed for both reducing disadvantageous behavior (for instance, gaming the system) and fostering advantageous behavior, like reflecting, abstracting and elaborating. The goals of scaffolding are helping the learner to solve the task at hand, but also increasing the understanding to support future task solving (Reiser, 2004). Scaffolding can come in two forms (Reiser, 2004): making the task easier (structuring) vs. making the task harder (problematizing concepts). Fourth graders working in dyads with a tablet app benefitted from prompts that problematized concepts by requesting explanations in that they had a higher discussion quality, but they also displayed more negative emotions and more off-task behavior (Schmitt & Weinberger, 2017). Seventh graders were found to perform better in a physics game running on tablets when they worked with peers compared to working alone (Chen & Law, 2016). This positive effect was even strengthened when question prompts were provided. However, there are also indications that both – collaborative condition and question prompts - had a negative impact on self-reported motivation (Chen & Law, 2016), pointing again at the need of challenging learners to increase their opportunities to learn, while keeping their motivation flow high (see also, Deater-Deckard, El Mallah, Chang, Evans, & Norton, 2014).

Fostering regulation of interactions

While structuring methods are usually set before the learning activities start, fostering learners' regulation can occur during the ongoing learning activities. One prominent way of regulation are awareness tools that visualize qualities of learners' interactions whilst they take place, e.g., showing how much learners are participating (Gijlers et al., 2013). On one hand, raising awareness showed to positively influence variables like transactivity and shared task focus (Gijlers et al., 2013), cohesion, positive attitude towards collaborative learning, and reduced conflicts (Phielix, Prins, & Kirschner, 2010), or fostered coordination of the collaborative learning situation (Janssen, Erkens, Kanselaar, & Jaspers, 2007). On the other hand, the overall influence of awareness tools remains limited as the aforementioned studies could not find a general effect of an awareness tool on the quality of the group product (Phielix et al., 2010) or equality of participation (Janssen et al., 2007). It appears to be a crucial factor to what extent learners are motivated and able to effectively use an awareness tool to alter their behavior patterns in order to foster positive group interactions and learning (Janssen, Erkens, & Kirschner, 2011). This goal may be achieved by scripting the use of the awareness tool (Janssen et al., 2011; Tsovaltzi et al., 2015).

Conclusion

In this chapter, two major branches of computer-supported collaborative learning, namely computer-mediated vs. co-present collaborative learning, were discussed. Computer-mediated collaborative learning scenarios encompass asynchronous as well as synchronous learning activities. The learning scenario might be reduced to a text-based forum or simulate a richer social situation, like in virtual worlds or video-conferences. The scenario might foresee one-to-one communication or address large amounts of users at the same time, as in the case of MOOCs. Learning anywhere and anytime at one's own pace is one of the benefits of such learning environments. Typical issues in computer-mediated collaborative learning are the lack of social context cues, free riding, a greater need for explicit coordination, higher attrition and large drop-out rates. In co-present collaborative learning scenarios, learners are physically together and engage in synchronous learning activities. The scenario might foresee to share one device, which is particularly suited for tabletops and tablets, or to provide every learner with an own device, for example a hand-held. Co-present scenarios can take into account the growing recognition of the role of the body in learning and afford hands-on embodied learning experiences. Typical issues in co-present collaborative learning are how learners divide the available working space when the device is shared, how to prevent individual members of exclusively controlling input and dominating the group as well as how to sensibly combine embodied learning experiences with phases of reflection (Danish et al., 2015; Schmitt & Weinberger, 2017).

Scripting interactions or providing awareness tools showed to be promising means in supporting CSCL. Next to the positive effects, scripts can come with negative side effects, such as over-scripting or demotivation (see e.g., Chen and Law, 2016; Dillenbourg, 2002;

Schmitt & Weinberger, 2017), and awareness tools' effects might fall short of expectations. Hence, more research is needed in refining instructional means to further develop CSCL practices in primary and secondary education. Another emerging topic concerns the question how to increase salience of feedback through awareness tools and leverage their effectiveness in CSCL environments. To that end, for example scripting the use of awareness tools has been proposed (Janssen et al., 2011; Tsovaltzi et al., 2015).

Beyond designing for productive, but isolated CSCL scenarios, orchestration of collaborative with individual learning scenarios, enriched with various types of computer-support will allow learners to encounter diversified challenges, finally leading to progressively expanding knowledge and skills. While CSCL practice typically is embedded in larger learning arrangements, research on how to productively coordinate scaffolding across learning experiences occurring at different social levels is still scarce (Prieto, Dimitriadis, Asensio-Pérez, & Looi, 2015). Future research will need to focus on the conditions of productive orchestration of learning scenarios and their support, for example awareness tools and scaffolding, with the ultimate goal of fully developing CSCL's potential for acquiring both, domain knowledge and key-competencies. Beyond sequencing learning arrangements, the question expands to how learners use different devices and link and document their learning activities in a larger learning landscape. Such a learning landscape bridges formal and informal learning in schools, social networks, any other place and time using mobile devices.

References

Abrahamson, D. (2017). Embodiment and mathematics learning. In K. Peppler (Ed.), *The SAGE encyclopedia of out-of-school learning* (pp. 247-252). Thousand Oaks, CA: SAGE.

Bechky, B. A. (2003). Sharing meaning across occupational communities: The transformation of understanding on a production floor. *Organization Science*, *14*(3), 312–330. doi:10.1287/orsc.14.3.312.15162

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., & Gijselaers, W. H. (2006). Common ground, complex problems and decision making. *Group Decision and Negotiation*, *15*(6), 529–556. doi:10.1007/s10726-006-9030-1

Chen, C. H., & Law, V. (2016). Scaffolding individual and collaborative game-based learning in learning performance and intrinsic motivation. *Computers in Human Behavior*, *55*, 1201–1212. doi:10.1016/j.chb.2015.03.010

Cohen, E. G., & Lotan, R. A. (1995). Producing equal-status interaction in the heterogeneous classroom. *American Educational Research Journal*, *32*(1), 99–120. doi:10.3102/00028312032001099

Crook, C. (1998). Children as computer users: The case of collaborative learning. *Computers & Education*, *30*(3/4), 237–247.

Danish, J. A., Enyedy, N., Saleh, A., Lee, C., & Andrade, A. (2015). Science through technology enhanced play: Designing to support reflection through play and embodiment. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015: Vol. 1* (pp. 332–339). Gothenburg, Sweden: The International Society of the Learning Sciences.

Deater-Deckard, K., El Mallah, S., Chang, M., Evans, M. A., & Norton, A. (2014). Student behavioral engagement during mathematics educational video game instruction with 11-14 year olds. *International Journal of Child-Computer Interaction*, *2*(3), 101–108. doi:10.1016/j.ijcci.2014.08.001

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL?* (pp. 61–91). Heerlen, Netherlands: Open Universiteit Nederland.

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

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning: from design to orchestration. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-Enhanced Learning: Principles and Products* (pp. 3–19). New York, NY: Springer Netherlands.

Ertl, B., Fischer, F., & Mandl, H. (2006). Conceptual and socio-cognitive support for collaborative learning in videoconferencing environments. *Computers & Education, 47*(3), 298-315.

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

Gelmini-Hornsby, G., Ainsworth, S., & O'Malley, C. (2011). Guided reciprocal questioning to support children's collaborative storytelling. *International Journal of Computer-Supported Collaborative Learning, 6*(4), 577–600. doi:10.1007/s11412-011-9129-5

Gijlers, H., Weinberger, A., van Dijk, A. M., Bollen, L., & van Joolingen, W. (2013). Collaborative drawing on a shared digital canvas in elementary science education: The effects of script and task awareness support. *International Journal of Computer-Supported Collaborative Learning*, *8*(4), 427–453. doi:10.1007/s11412-013-9180-5

Hornecker, E., Marshall, P., Dalton, N., & Rogers, Y. (2008). Collaboration and interference: Awareness with mice or touch input. *Proceedings of the Conference on Computer Supported Cooperative Work.* San Diego, USA: ACM Press.

Hsiao, I. Y. T., Yang, S. J. H., & Chu, C. J. (2015). The effects of collaborative models in second life on French learning. *Educational Technology Research and Development, 63*(5), 645-670. doi:10.1007/s11423-015-9379-4

Janssen, J., Erkens, G., Kanselaar, G., & Jaspers, J. (2007). Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education*, *49*(4), 1037–1065. doi:10.1016/j.compedu.2006.01.004

Janssen, J., Erkens, G., & Kirschner, P. A. (2011). Group awareness tools: It's what you do with it that matters. *Computers in Human Behavior*, *27*(3), 1046–1058. doi:10.1016/j.chb.2010.06.002

Janssen, J., Erkens, G., Kirschner, P. A., & Kanselaar, G. (2010). Effects of representational guidance during computer-supported collaborative learning. *Instructional Science*, *38*(1), 59–88. doi:10.1007/s11251-008-9078-1

Jordan, K. (2015). Massive Open Online Course completion rates revisited: Assessment, length and attrition. *International Review of Research in Open and Distance Learning, 16*(3), 341-358. doi:10.19173/irrodl.v16i3.2112

Kerr, N. & Bruun, S. (1983). Dispensability of member effort and group motivation losses: Free-rider effects. *Journal of Personality and Social Psychology, 44*(1), 78-94.

Kreijns, K., Kirschner, P. A., & Vermeulen, M. (2013). Social aspects of CSCL environments: A research framework. *Educational Psychologist, 48*(4), 229-242. doi:10.1080/00461520.2012.750225

Lampe, C., Vitak, J., Gray, R., & Ellison, N. (2012). Perceptions of Facebook's value as an information source. *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 3195–3204). ACM. doi:10.1145/2207676.2208739

Ligorio, M. B., & Van der Meijden, H. (2008). Teacher guidelines for cross-national virtual communities in primary education. *Journal of Computer Assisted Learning*, *24*(1), 11-25.

Manlove, S., Lazonder, A. W., & de Jong, T. (2009). Trends and issues of regulative support use during inquiry learning: Patterns from three studies. *Computers in Human Behavior, 25*(4), 795–803. doi:10.1016/j.chb.2008.07.010

Mercier, E. M., Higgins, S. E., & da Costa, L. (2014). Different leaders: Emergent organizational and intellectual leadership in children's collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning*, *9*(4), 397–432. doi:10.1007/s11412-014-9201-z

Moed, A., Otto, O., Pal, J., Singh, U. P., Kam, M., & Toyama, K. (2009). Reducing dominance in multiple-mouse learning activities. In *CSCL 2009* (pp. 360–364).

Nihalani, P. K., Wilson, H. E., Thomas, G., & Robinson, D. H. (2010). What determines high-and low-performing groups? The Superstar Effect. *Journal of Advanced Academics, 21*(3), 500-529. doi:10.1177/1932202X1002100306

Nistor, N., & Neubauer, K. (2010). From participation to dropout: Quantitative participation patterns in online university courses. *Computers & Education, 55*(2), 663-672. doi:10.1016/j.compedu.2010.02.026

Noroozi, O., Weinberger, A., Biemans, H. J. A., Mulder, M., & Chizari, M. (2013). Facilitating argumentative knowledge construction through a transactive discussion script in CSCL. *Computers & Education, 61*, 59–76.

Park, J. H., & Choi, H. J. (2009). Factors influencing adult learners' decision to drop out or persist in online learning. *Educational Technology & Society, 12*(4), 207-217.

Phielix, C., Prins, F. J., & Kirschner, P. A. (2010). Awareness of group performance in a CSCL-environment: Effects of peer feedback and reflection. *Computers in Human Behavior*, *26*(2), 151–161. doi:10.1016/j.chb.2009.10.011

Puhl, T., Tsovaltzi, D., and Weinberger, A. (2015). A long-term view on learning to argue in Facebook: The effects of group awareness tools and argumentation scripts. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015: Vol. 1* (pp. 110-117). Gothenburg, Sweden: The International Society of the Learning Sciences.

Prieto, L. P., Dimitriadis, Y., Asensio-Pérez, J. I., & Looi, C.-K. (2015). Orchestration in learning technology research: evaluation of a conceptual framework. *Research in Learning Technology, 23*(1), 28019.

Raes, A., Schellens, T., De Wever, B., & Benoit, D. F. (2016). Promoting metacognitive regulation through collaborative problem solving on the web: When scripting does not work. *Computers in Human Behavior*, *58*, 325–342. doi:10.1016/j.chb.2015.12.064

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, *13*(3), 273–304. doi:10.1207/s15327809jls1303

Rick, J., Kopp, D., Schmitt, L., & Weinberger, A. (2015). Tarzan and Jane Share an iPad. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Vol. 1* (pp. 356–363). Gothenburg, Sweden: The International Society of the Learning Sciences.

Roschelle, J., Rafanan, K., Bhanot, R., Estrella, G., Penuel, B., Nussbaum, M., & Claro, S. (2010). Scaffolding group explanation and feedback with handheld technology: Impact on students' mathematics learning. *Educational Technology Research and Development*, *58*(4), 399–419. doi:10.1007/s11423-009-9142-9

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer Supported Collaborative Learning* (pp. 69–97). Berlin, Germany: Springer. doi:10.1145/130893.952914

Scheuer, O., Loll, F., Pinkwart, N., & McLaren, B. M. (2010). Computer-supported argumentation: A review of the state of the art. *International Journal of Computer-Supported Collaborative Learning*, *5*(1), 43–102. http://doi.org/10.1007/s11412-009-9080-x

Schmitt, L., & Weinberger, A. (2017). Collaborative Learning on Multi-Touch Interfaces: Scaffolding Elementary School Students. In B. K. Smith, M. Borge, E. Mercier, & K. Y. Lim (Eds.). *Making a Difference: Prioritizing Equity and Access in CSCL: The 12th International Conference on Computer Supported Collaborative Learning, Volume 1* (pp. 9-16). Philadelphia, PA, USA: The International Society of the Learning Sciences.

Schneps, M. H., Ruel, J., Sonnert, G., Dussault, M., Griffin, M., & Sadler, P. M. (2014). Conceptualizing astronomical scale: Virtual simulations on handheld tablet computers reverse misconceptions. *Computers & Education, 70*, 269–280.

Selwyn, N. (2009). Faceworking: Exploring students' education-related use of Facebook. *Learning, Media and Technology, 34*(2), 157-174. doi:10.1080/17439880902923622

Sitzmann, T., Ely, K., Bell, B. S., & Bauer, K. N. (2010). The effects of technical difficulties on learning and attrition during online training. *Journal of Experimental Psychology-Applied, 16*(3), 281-292. doi:10.1037/a0019968

Stahl, G. (2006). Supporting group cognition in an online math community: A cognitive tool for small-group referencing in text chat. *Journal of Educational Computing Research*, *35*(2), 103–122. doi:10.2190/Q435-7611-2561-720P

Stahl, G. (2013). Transactive discourse in CSCL. *International Journal of Computer-Supported Collaborative Learning*, *8*(2), 145–147. http://doi.org/10.1007/s11412-013-9171-6

Stanton, D., Neale, H., & Bayon, V. (2002). Interfaces to support children's co-present collaboration: multiple mice and tangible technologies. In G. Stahl (Ed.), *Computer support for collaborative learning: foundations for a CSCL community* (pp. 583-584). Mahwah, NJ: Lawrence Erlbaum Associates.

Tsovaltzi, D., Judele, R., Puhl, T., & Weinberger, A. (2015). Scripts, individual preparation and group awareness support in the service of learning in Facebook: How does CSCL compare to social networking sites? *Computers in Human Behavior, 53*, 577-592. doi:10.1016/j.chb.2015.04.067

Vanderlinde, R., Aesaert, K., & Van Braak, J. (2014). Institutionalised ICT use in primary education: A multilevel analysis. *Computers & Education*, *72*, 1–10. doi:10.1016/j.compedu.2013.10.007

Van Der Meijden, H., & Veenman, S. (2005). Face-to-face versus computer-mediated communication in a primary school setting. *Computers in Human Behavior, 21*(5), 831–859. doi:10.1016/j.chb.2003.10.005

Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research, 23*(1), 3-43.

Walther, J. B. (2011). Theories of computer-mediated communication and interpersonal relations. In M. L. Knapp & J. A. Daly (Eds.), *The handbook of interpersonal communication* (pp. 443–479). Thousand Oaks, CA: Sage.

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

Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning and Instruction*, *17*(4), 416–426. doi:10.1016/j.learninstruc.2007.03.007

Wodzicki, K., Schwämmlein, E., & Moskaliuk, J. (2012). "Actually, I wanted to learn": Study-related knowledge exchange on social networking sites. *Internet and Higher Education, 15*(1), 9-14. doi:10.1016/j.iheduc.2011.05.008

Zahn, C., Krauskopf, K., Hesse, F. W., & Pea, R. (2012). How to improve collaborative learning with video tools in the classroom? Social vs. cognitive guidance for student teams. *International Journal of Computer-Supported Collaborative Learning, 7*(2), 259–284. doi:10.1007/s11412-012-9145-0

Study II:

Schmitt, L. J., Rick, J., & Weinberger, A. (2019). BiCo: A bipolar continuous rating scale for children's technology evaluation. *Technology, Pedagogy and Education, 28*(5), 503–516. doi:10.1080/1475939X.2019.1661279.

The Version of Record of this manuscript has been published and is available in Technology, Pedagogy and Education (2019),
https://www.tandfonline.com/doi/10.1080/1475939X.2019.1661279

BiCo: A Bipolar Continuous Rating Scale for Children's Technology Evaluation

Lara Johanna Schmitt, Jochen Rick, and Armin Weinberger

Abstract

One challenge of child-computer interaction research is surveying variation of children's attitudes towards novel educational technology, which often results in an opinion ceiling effect. In this article, we introduce BiCo – a bipolar continuous rating scale, which builds on children's ability to draw relative comparisons. We elaborate on the development of BiCo and how we designed it to better survey children's attitudes. Beyond addressing the opinion ceiling effect, BiCo is suitable for surveying a wide range of concepts and its invertible design enables simulation of inverted items. We provide data comparing BiCo vs. the widely used Smileyometer instrument when used by fourth graders (around 10 years old), demonstrating that BiCo mitigates the opinion ceiling effect. We discuss BiCo as a new tool for children's technology evaluation and provide directions for future research.

Keywords: opinion ceiling effect, questionnaire, Smileyometer, BiCo, child-computer interaction

1. Empirical Research with Children

Developing sound educational technologies for children requires research with young learners as the primary target group. Knowing what makes educational technologies effective and enjoyable learning tools is a key foundation for the successful integration of educational technologies into the classroom and other learning settings. Therefore, it is important to investigate and further develop research methods to meaningfully capture learning processes and children's perceptions in interacting with technologies. There are established methods for involving children in actively designing and evaluating technology applications (Fails, Guha, & Druin, 2012). Many of these focus on the formative design process, such as using children as informants or design partners. For example, Chimbo and Gelderblom (2018) successfully involved children and teenagers in a participatory design process to develop a tutoring application prototype. To utilize these methods, researchers / developers must usually work closely with the young participants. Due to the effort required, there is an inherent trade-off between the number of participants and the number of iterations. Hence, it is recommended to keep to less than ten participants; this small sample will generate enough feedback to inform the next iteration (Nielsen, 1993).

For efficiently assessing attitudes and testing hypotheses with a larger sample size, however, reliable quantitative measurements by questionnaires are required (Hall, Hume, & Tazzyman, 2016). As Read and Markopoulos (2013) argue, empirical work is still an open challenge for child-computer interaction (CCI); part of that challenge is 'the development of robust methods' (p. 4). In particular, CCI research often values criteria such as usability and fun that are inherently hard to measure through observation (Read, 2008). While it may be possible to improve observation methods, such as utilizing eye tracking to study focus (Fowler, 2013), or analysis of interactions through video and log file data (Hall et al., 2016), participants' self-reporting is a valuable complementary source of information (Sim, MacFarlane, & Read, 2006). Yet, surveying children is not trivial. Traditionally, researchers have been sceptical about collecting information directly from children because it was assumed that the still developing cognitive abilities and literacy / language skills would not allow for reliable and valid responses (Bell, 2007; Fuchs, 2005; Scott, 2000); surveys focused more on teachers' or parents' indirect opinions about children, although those reports, too, are likely to be biased (Beitchman & Corradini, 1988; Scott, Brynin, & Smith, 1995). As children can allow for direct insights into their own world, it is important and also possible to develop methods to do research *with* instead of *about* children (Beitchman & Corradini, 1988; Fuchs, 2005; Scott, 2000). Children can be competent and active participants in research, if researchers do not treat them as underdeveloped adults but are willing to understand their view of the world and to speak their language (Fraser, Flewitt, & Hammersley, 2014). For example, it is critically important to carefully craft the questionnaire to be child friendly (Colombo & Landoni, 2014). Early work suggests that children can successfully respond to questionnaire items when they are child friendly: e.g., Scott et al. (1995) administered a questionnaire with seven smiley phases (from sad over neutral to happy) that enabled children aged 11-16 to rate their mood.

There is a continuous interest in developing instruments that are particularly accessible to children (Bell, 2007; Horton, 2013). In this article, we introduce the new 'BiCo' instrument we developed to better capture children's technology evaluation. We will introduce the topic of surveying children and present commonly used instruments to capture children's perceptions of technologies along with these instruments' advantages and disadvantages. We will substantiate how and why we arrived at the design of BiCo as a new instrument for children's technology evaluation. Last, we will provide first data comparing BiCo to the established 'Smileyometer' instrument as a 'proof of concept' in the context of a computer-supported learning study.

2. Questionnaires for Children

Questionnaires are used in empirical research to assess performance, personality, or attitudes. The challenges in designing questionnaires for children are similar to those when designing for adults, but accentuated, because children's cognitive and social skills are still under development (Borgers, de Leeuw, & Hox, 2000). Although surveying children is suitable from age 8 years on (Borgers et al., 2000), ambiguity and complex questions are difficult to understand for children. In a study by de Leeuw and Otter (1995), reliability of questionnaire responses was reduced for younger children (9 years old) compared to older children (14 years old), especially when the questions were ambiguous; both age groups showed worse reliability with complex compared to simple questions. Negative questions, e.g., 'the game was not good', are commonly not interpreted correctly (Marsh, 1986). De Leeuw and Otter (1995) advise to keep the items of a questionnaire for children very clear and simple. Understanding items and accurately rating their own feelings or beliefs can be challenging for children (Beitchman & Corradini, 1988). Long and boring questionnaires can result in reduced motivation and concentration, leading to a satisficing answering strategy of giving quick and superficial answers. When surveys are conducted in schools, proximity of classmates can influence responses, as children tend to answer in a way that would please their peer group (Scott, 2000).

Currently, the existing instruments to survey children reveal reliability problems. For example, Horton, Read and Sim (2011) asked 19 children aged 6–7 to indicate which technologies they have at home using a pictorial questionnaire. One week later, they asked the same question in a verbal format. Though it was a straightforward question and the time between responses was small, none of the children stated the same technologies in both questionnaires. One explanation could be children's tendency to interpret questions literally (e.g., Scott et al., 1995): if the device in the picture was not exactly the one the children own, they would indicate not to possess it even though they own a similar device (Horton et al., 2011).

When building a questionnaire, the initial decision to be taken is whether the items should have an open or a closed format (Moosbrugger & Kelava, 2008). Open items, e.g., completing a sentence or writing a short essay, allow for free answers and are useful for capturing concepts like creativity that would be hard to survey otherwise. They can also limit the usefulness of guessing strategies in a performance test. In empirical research, open answers are typically coded / quantified for further analysis; coding can require considerable effort to ensure objectivity. Also, non-response behavior is often a significant problem when surveying children with open items (Smith & Platt, 2013). Closed items, e.g., choosing one answer of several choices or sorting concepts into categories, confine the answers to a limited selection. In comparison to open items, the response rates of closed items are higher (Smith & Platt, 2013) and they can be analyzed efficiently. A particularly common and useful class of closed questionnaire items are rating scales.

2.1 Rating Scales and the Smileyometer

A rating scale is an instrument to measure subjective judgments along a preset axis. Rating scales are efficient. Once a participant becomes proficient with a scale, they can quickly answer many questions using that scale. For the researcher, quantifying the responses is easy. Because numerical values are suitable for statistical analysis, rating scales are common in empirical research. Rating scales can be classified along several dimensions (Moosbrugger & Kelava, 2008).

First, they can be either continuous or discrete. Discrete scales consist of two or more answering options. The maximum number of answering options is theoretically open, but seven options are considered to be an adequate limit, as people do not often use the full range of options provided (Moosbrugger & Kelava, 2008). Continuous scales do not have discrete answering options; the response is given freely along a labeled axis. They allow for especially fine-grained judgments. Second, there are unipolar and bipolar rating scales; the use of a unipolar vs bipolar format is dependent on the concepts to be measured (Moosbrugger & Kelava, 2008). Unipolar rating scales measure preferences along one dimension, such as agreement (e.g., from 'don't agree' to 'strongly agree'). Bipolar rating scales measure preferences along two (usually opposed) dimensions, such as disagreement and agreement (e.g., from 'strongly disagree' to 'strongly agree'). If a response behavior into a certain direction is expected, for example when respondents are expected to rate something rather positively, it can be useful to use an asymmetric format favoring one pole to allow for a higher differentiation in the expected rating range (Moosbrugger & Kelava, 2008). Third, rating scales can differ in denomination: numerical, verbal or symbolic (Moosbrugger & Kelava, 2008). These formats can also be combined. In the case of a numerical format, the scales are labeled with numbers. This can be problematic since there may be a discrepancy between the numbers and the actual distances between them as perceived by the individual participants. This phenomenon plays a smaller role in the case of the verbal format: Here, the scales are labeled with words and are therefore supposed to evoke more consistent interpretations by the participants. The symbolic format uses neither numbers nor words, but a symbolic display of the preference response (e.g., '− −, − +, + +') and thusly tries to avoid subjective interpretations of numbers or words (Moosbrugger & Kelava, 2008).

Good questionnaire design tries to minimize the number of scales used as each additional instrument imposes a cognitive burden. When surveying adults, it is typical to limit the questionnaire to one or two versions of the Likert scale to measure different concepts. For instance, a 5-point Likert scale of 'strongly disagree, disagree, neutral, agree, strongly agree' might be used to assess agreement, while a 4-point Likert scale of 'often, sometimes, seldom, never' might assess frequency. While children can use verbal scales, decoding words provides a higher cognitive burden than for adults. Recently, there has been interest in providing symbolic scales to ease that burden. While these might have written labels, it is easier for the children to associate meaning with the pictorial representation. For instance, the Thumbs-Up Scale (Kano, Horton, & Read, 2010) provides thumbs up / down analogues to sample proficiency.

A widely used symbolic scale designed for children is Smileyometer (Read, 2008), a scale designed to measure fun, a particularly important concept in CCI. As demonstrated in Figure 1, it employs five smiley faces to indicate attitudes ranging from 'awful' to 'brilliant'. Often, Smileyometer is applied twice, before and after a technology intervention to measure children's expectations and how much they were satisfied or exceeded (Sim et al., 2006). Though it has been used successfully, a documented deficiency of this scale is the opinion ceiling effect.

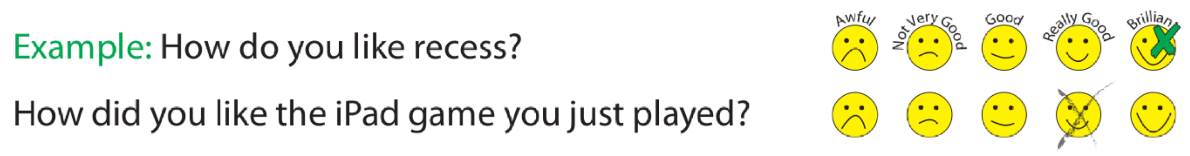


Figure 1. The Smileyometer rating scale with an example and one survey item

2.2. The Opinion Ceiling Effect

Appropriate item difficulty is critical to questionnaire design. While item difficulty is obvious in achievement tests with items that have a correct answer, the concept extends to questionnaires, where the term refers to the tendency to agree to an item (Moosbrugger & Kelava, 2008). The higher the difficulty, the more people tend to agree to an item. Items should reveal actual variance to differentiate between people with different degrees of a certain characteristic. Hence, questions should hold a medium difficulty index, i.e., differentiating to a maximal extent between people with high and low feature characteristic, with variability across questions and respondents. Single items with a higher or lower difficulty can add value as they differentiate in the more extreme areas of the feature characteristic.

The opinion ceiling effect occurs when responses are so positive for all of the items and conditions that it is hard to detect any potentially existing difference between children. Children's overly positive responses to a technology experience are a well-documented phenomenon in CCI research in general and with Smileyometer in particular. For instance, when sampling 8–9 year olds about a robot with Smileyometer, some conditions received over half 'brilliant' (the most positive) response (Looije, Neerincx, & de Lange, 2008). Similarly, children aged 7–8 rated science learning software as 'brilliant' over half the time (Sim et al., 2006). These findings are not uncommon in CCI research; often, the mean for

Smileyometer questions is between 4 and 5 (with 5 being the most positive option), see e.g. Sim, Horton, and McKnight (2016), or Johnson, Shum, Rogers, and Marquardt (2016).

There are several contributing factors to the opinion ceiling effect in CCI work.

(1) The technology being evaluated is novel and matured by the time it gets to empirical research, so that we expect children to have a positive opinion of it. For example, in one of our studies about children's collaborative learning with a tablet app, the children showed more positive than negative emotions, and worked diligently with the app throughout forty minutes of intervention time (Schmitt & Weinberger, 2019).

(2) Younger children are inclined to give extreme responses. When asked to answer questions where a middle value was the expected answer based on the information presented, there was a significant correlation of age and answering either extremely positive or negative: age 5–6 >7–9 >10–12 (Chambers & Johnston, 2002). That effect was amplified for emotional questions. While extreme responses in that pediatric psychology study were both positive and negative, extreme positive is more likely in CCI research. To study the effect of age, Read and MacFarlane (2006) conducted a simple questionnaire of video games, comparing a group of younger children (7–9) with a group of older children (12–13). There was a significant difference between the two groups with younger children tending towards higher ratings. Over 40% of the younger children, but less than 10% percent of the older children rated every game they reviewed with the highest rating.

(3) Children might display an acquiescence bias, feeling that positive responses are expected of them (e.g., Colombo & Landoni, 2014). A wish to please the adult/researcher can lead to compliance (Bell, 2007). While this is related to the first factor (children may feel the need to rate a mature software positively), it is influenced by the relationship between the researchers and the participants. For instance, if one researcher plays a 'Wizard of Oz' game with a child, then the child might evaluate the game positively to please that researcher.

(4) Children might display a satisficing effect, choosing to simply answer all the questions the same. It is easier to answer all questions positively than to really engage with each one. That factor is particularly likely for a long, tedious questionnaire (Borgers et al., 2000). Compared to completing an enjoyable CCI task, filling out a questionnaire is less interesting. In general, questionnaires designed for CCI research

should be kept short (Law, Watkins, Barwick, & Kirk, 2013; Read & MacFarlane, 2006).

In addition to these four factors, details of the technology, the children and the study setting may impact the occurrence of an opinion ceiling effect in a specific research study. In spite of the challenges inherent in surveying children, we think that innovation can mitigate the opinion ceiling effect. In several ways, Smileyometer is already a useful instrument in that regard. First, it is easy for children to parse and requires less cognitive effort than a verbal scale. Second, as it is more enjoyable to use, it might also minimize satisficing strategies. Third, it addresses the positive inclination by providing an asymmetric scale: instead of being neutral, the middle value ('good') is positive. Hall et al. (2016) found the positivity of smiley face Likert scales to be a factor in achieving variance in children's responses. In their study, they improved the quality of answers by providing the children with an all positive version of a smiley Likert scale ('Five Degrees of Happiness'), i.e. removing neutral or sad faces completely. In contrast, a common way to prevent the acquiescence bias is asking negative questions (Moosbrugger & Kelava, 2008). That strategy is not possible with Smileyometer. Considering that children have difficulties responding accurately to negative questions (Bell, 2007; Marsh, 1986), that technique may not be suitable for children anyway.

2.3. An Alternative: Comparisons

Smileyometer was introduced as part of Fun Toolkit, which also introduced the 'Fun Sorter' and 'Again-Again' (Sim & Horton, 2012). Again-Again asks children whether they would want to use the technology again (yes, maybe, no). The Fun Sorter asks children to rank multiple technologies along various dimensions (e.g., fun). Another instrument that uses comparisons is 'This or That' (Zaman, Abeele, & De Grooff, 2013), where children have to indicate which game, out of two games they played, they favor, related to questions asking e.g. which one they would like to play again or they would like to take home. The Fun Sorter, Again-Again and This or That are suited to the simultaneous evaluation / comparison of multiple technologies or applications. Such comparisons are particularly accessible to children (Hanna, Neapolitan, & Risden, 2004). Children generally have an easier time with these comparisons, allowing researchers to draw conclusions such as which game is more fun on average. Using the Fun Sorter, children aged 7–8 were able to rank three different software products differently for the dimensions fun, ease of use and learning, demonstrating that children are able to separate these concerns (Sim et al., 2006). Comparison instruments are a useful way to avoid the opinion ceiling effect, particularly for younger children. Table 1 gives an overview on the different existing survey instruments and their advantages and disadvantages.

Table 1
Overview on existing survey instruments for children's technology evaluation.

<table>
<tr><th>Instrument</th><th>How it works</th><th>Advantages</th><th>Disadvantages</th></tr>
<tr><td>Smileyometer</td><td>Symbolic scale with 5 smiley faces from ‘awful’ to ‘brilliant’</td><td>- Enjoyable and easy to use
- Asymmetric scale</td><td>- Often associated with the opinion ceiling effect
- Cannot include inverted items
- Not well suitable for concepts beyond ‘fun’</td></tr>
<tr><td>Fun Sorter</td><td>Ranking of the multiple technologies/applications that the child has interacted with (different dimensions, e.g. fun)</td><td rowspan="3">- Drawing comparisons is accessible for children
- Avoiding the opinion ceiling effect</td><td rowspan="3">- Only suitable for the simultaneous evaluation of multiple technologies/applications (within-subjects design)</td></tr>
<tr><td>Again Again</td><td>Table in which the child indicates if it would want to use the technology/application again (‘yes’, ‘maybe’, ‘no’)</td></tr>
<tr><td>This or That</td><td>Child indicates which technology/application they favor (different questions, e.g. which one would they want to use again)</td></tr>
</table>

When employing a comparison instrument in a within-subjects design, each child needs to play every game to be able to choose which he or she prefers. To address order effects, e.g., participants tend to prefer the last game they played, the play order is counter-balanced across participants. The counter-balancing technique could also be employed across experimental conditions. However, the counter-balancing technique is often not feasible in empirical research, because significant order effects are expected: E.g., Harris et al. (2009) found that children who switched from a single-touch condition to a multi-touch condition worked well in both, whereas children who switched from multi-touch to single-touch had problems with the latter. Apparently, taking away a feature is more problematic than not having that feature to begin with. So, while children can use comparison, it is not always realistic to integrate multiple applications into empirical research. That may be one reason that Smileyometer is more frequently employed than the Fun Sorter. ‘Again-Again’ has recently also been adapted to be used in a between-subjects design (Sim & Cassidy, 2013; Sim et al., 2016), i.e. for the evaluation of only one technology, with questions like ‘If the game were free would you download it from the app store?’

3. Developing the BiCo Rating Scale

Based on the literature on surveying children and the advantages and disadvantages of existing instruments, we developed BiCo (figure 2). The basic underlying idea of the **Bi**polar **Co**ntinuous (BiCo) rating scale is that the children compare two statements to each other and mark their agreement with an X along a continuous line.

Developing BiCo was led by several design choices:

(1) Bipolar instead of unipolar: Inspired by children’s ability to make relative comparisons, we implemented a bipolar scale with two opposing statements on each side. These two bipolar statements should enable the children to draw a comparison between them and give more nuanced responses. Providing two statements to compare can also combat satisficing strategies of choosing the same answer in each question, because the side of the positive attitude statement can be switched within a series of items.

(2) Continuous instead of discrete: We opted for a continuous scale to allow for more fine-grained responses. Even someone who overwhelmingly agreed with one statement over the other might back off a little bit from indicating ‘only this’ to indicate slight agreement with the other position. This would differentiate that participant from someone who agreed entirely with the statement. Additionally, most statistics-based tests assume a continuum of response; discrete scales (e.g., Likert scales) just approximate that.

(3) Verbal denomination: Instead of numerical or symbolic, we chose the verbal format. Verbal formats are supposed to evoke consistent interpretations by the participants

and the verbal format lent itself well to the bipolar presentation of two opposing statements.

We created BiCo equivalents for Smileyometer questions (Table 2) to be administered in a post-questionnaire of a study in the context of computer-supported collaborative learning (see section 4). Sometimes it made sense to create diametrically opposed positions (e.g., Table 2, item f). At other times, it made more sense to create choices that make the decision harder. Our goal was to create meaningful comparisons that conceptually addressed the same subject we sampled with the Smileyometer questions rather than create morphological equivalents. For instance, for item d, we judged it to be too easy for children to agree with the sentiment ‘My partner was helpful’ over ‘My partner was unhelpful’ (i.e., acquiescence bias). So, we chose ‘My partner distracted me’, which could be simultaneously true with ‘My partner was helpful’, forcing children to weigh the relative merits of the statements. Given children’s problems with negatively stated items (Borgers, Sikkel, & Hox, 2004), we also kept the statements positive. For instance, when converting question e, we chose BiCo equivalents of "we worked well together" and "we would have been better off working alone." For morphological equivalence, the latter should have been "we did not work well together;" however, we wanted to sample how they valued collaboration. For that, creating a tension between the two statements, both of which could be true to some degree, made more sense. In this case, the value of collaborating as the underlying concept to be measured was emphasized rather than the issue of how well they worked with each other. For the majority of questions (a, b, c & f), the BiCo equivalents matched closely to the Smileyometer wording. As the children in our study spoke German, all questions and instruments are in German; here we present them in English. The German versions were iteratively revised: We discussed and refined the wording within our research team, as well as with an expert on media usage in primary schools from another institution. Iterations are necessary to formulate the items to be usable by children and to avoid ambiguity, i.e. the items are well understandable and fit the children’s reading level. The English translations are intended only for our adult readers. Similar language polishing may be necessary to utilize these exact questions with English-speaking children.

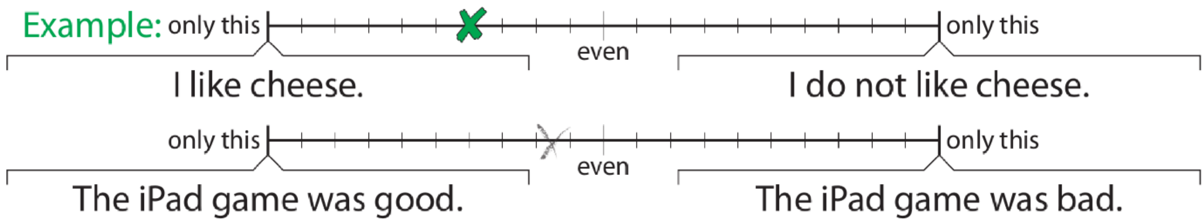


Figure 2. The BiCo rating scale with an example and one survey item

Table 2.
Six items presented in Smileyometer vs. BiCo format

#	Smileyometer Questions	BiCo Statements to Compare	
(a)	Would you like to continue?	I would like to stop.	I would like to continue.
(b)	Did you learn something about math with the iPad?	I learned a lot using the iPad.	I did not learn anything using the iPad.
(c)	Did you help your partner a lot?	I helped my partner a lot.	I did not help my partner.
(d)	How helpful was your partner?	My partner distracted me.	My partner was helpful.
(e)	How well did you work together?	We worked well together.	We would have been better off working alone.
(f)	How did you like the iPad game you just played?	The iPad game was bad.	The iPad game was good.

Figures 1 & 2 show how we present the instruments to the children, beginning each instrument use with a filled in example (attitude towards recess and cheese respectively). To prevent visual clutter, Smileyometer's textual labels are only applied to the example. These figures demonstrate how the same topic (attitude towards a technology/application) can be sampled with each instrument and how children respond to each instrument.

BiCo offers several advantages over Smileyometer. First, Smileyometer always has the most positive response on the right side, so a satisficing strategy is feasible. With BiCo, the side of the positive attitude statement can be switched. This achieves a similar effect to presenting inverted items, which is a common strategy to avoid acquiescence bias in adults (Moosbrugger & Kelava, 2008). Rather than inverting the question, BiCo inverts the scale (e.g. compare items c and d, Table 2). Together with the approach of letting the participants draw relative comparisons, this is hypothesized to mitigate the opinion ceiling effect. Secondly, the scope of questions that can be asked with BiCo is larger. Smileyometer was designed to measure subjective fun on a scale from awful to brilliant (figure 1). While fun is important, it is not the only criterion to consider in child-computer interaction. While other instruments can be employed (e.g., the Thumbs-Up Scale to measure perceived skill levels; Kano et al., 2010), this forces children to use multiple instruments. One approach to avoid introducing a new instrument is to continue to use Smileyometer even if the match between the Smileyometer textual answers and the questions are not entirely in agreement. For instance, when assessing a robotic agent with Smileyometer, children were asked, ‘Would you like to talk some more with the robot / chatbot some time?’ (Looije et al., 2008); ‘not very good’ is not an entirely proper response to that question. We too employ that strategy to extend the scope of Smileyometer (e.g., item b ‘Did you learn something about math with the iPad?’ is not properly answered with ‘awful’). In these cases, the children probably interpret that the visual smileys demand an attitude response to the question. Thus, item a ‘Would you like to continue?’ can be interpreted as ‘How would you feel about continuing the game?’, a grammatically difficult question to parse. Still, some topics do not lend themselves to even an extended interpretation of Smileyometer. Utilizing BiCo’s broader scope, we are able to survey topics that are less of an ideal match for Smileyometer, e.g., whether iPads should be employed for mathematics class.

In sum, BiCo advantages are: With its bipolar format, it builds on children’s capability to draw comparisons and should therefore mitigate the opinion ceiling effect. Satisficing strategies are addressed by inverting the scale within a set of items. The continuous format allows for nuanced responses. The verbal denomination is supposed to evoke consistent interpretations and works well with the bipolar presentation of two opposing statements. BiCo is suitable in between-subjects design studies, where research participants evaluate only one experimental condition / technology without comparing it to another. Finally, BiCo enables to survey a wide range of concepts while sticking to the same instrument.

4. A study comparing BiCo and Smileyometer

4.1 Background: Encountering the opinion ceiling effect

We encountered the opinion ceiling effect in one of our studies about children’s collaborative learning with ‘Proportion’ (e.g., Rick, 2012; Schmitt & Weinberger, 2019). Proportion is an iPad application designed to foster proportional reasoning of children around the age of 10 years. In said study, we focused on the issues of collaboration and multi-touch; we had three conditions that were tested in a between-subjects design study: individuals, dyads working in multi-touch, and dyads working in single-touch. Multi-touch input means that the tablet is capable of sensing several touches simultaneously (by several fingers and / or several users); single-touch input means that only one touch at a time can be sensed. The study was done in three German primary schools with 125 fourth graders (aged 9–12, M=10 years and 3 months, SD=6.5 months). Each session took 90 minutes. First, participants were randomly assigned to a group / condition. Next, they completed a short questionnaire and mathematics test. Then, they spent 40 minutes working with the iPad. Finally, participants completed a short questionnaire and test. In the pre-questionnaire, we gathered socio-demographic data as well as attitudes towards school, math, individual and collaborative learning. The post-questionnaire (see table 2) asked about the attitudes towards the app (items a and f), subjective learning gain (item b) and, only in the collaborative conditions, aspects of the collaboration (items c, d and e). The questionnaires employed the Smileyometer scale (except for the socio-demographic items); see figure 1 for an example of Smileyometer as used in the study. To reduce the acquiescence bias in Proportion, we took a hands-off approach. While we thanked the children for their participation, we did not inform them that we had developed the app, nor did we help them with the game.

After collecting at the first school (n=43), we analyzed the data. While Smileyometer worked well in the pre-questionnaire with mean values ranging from 3.21 to 4.33 (1=awful... 5=brilliant), the post-questionnaire revealed strong opinion ceiling effects: Children tended to choose ‘brilliant’, with the mean values ranging from 4.29 to 4.67, producing only little variance. While the median response ranged between 3 and 5 in the pre-questionnaire, it was 5 for every question in the post-questionnaire, meaning that over half of answers were ‘brilliant’ for every question. Also, no participant answered all questions in the pre-questionnaire with the most positive option, but 16 of the participants (37%) did so in the post-questionnaire. While Smileyometer in general can be a useful instrument, it did not appropriately differentiate in our post-questionnaire.

4.2 Methods: BiCo vs. Smileyometer

We developed and utilized BiCo as a possible solution and employed it, along with

Smileyometer, in the second and third school; see figure 2 for an example of BiCo as used in the study. To compare both instruments with each other, we utilized a within-subjects design, i.e. aggregating the experimental factors, and the data of school 2 and 3.

When collecting data at the second and third school (n=82), the original questionnaires were extended with the BiCo questions. In our post-questionnaire, we arranged the BiCo statements so that our expected response was randomly switching sides (see table 2) to address the acquiescence bias. We presented both Smileyometer and BiCo to each student in order to compare BiCo and Smileyometer to assess BiCo's utility as an instrument to survey children's opinion on technology. The mean age of the sample was 10 years and 5 months (SD=6.6 months). n=36 students were female (44%). The following results compare the two formats of Smileyometer vs. BiCo of the six items of the post-questionnaire in a within-subjects-design.

4.3 Results: BiCo vs Smileyometer

A common way to establish the validity of a new instrument, such as BiCo, is to compare its results to an existing validated one, like Smileyometer. To do that, we normalized the data. For BiCo, we measured the distance from the negative pole, i.e., the unexpected response, to the response and divided it by the length of the scale, resulting in a value between 0 and 1. Smileyometer responses were similarly converted to the range 0–1 (0, 0.25, 0.50, 0.75, 1) from 'awful' to 'brilliant'. When a child marked more than one smiley, we averaged the result. Initially, a few children interpreted BiCo as demanding two Xs, one for each statement away from even. For those we added the distance from the negative pole to the first X and the distance from even to the second X to achieve one value. This behavior occurred in the first two experiment sessions with BiCo (n=26) for 15% of the given answers. In the following five experiment sessions (n=56), we verbally emphasized that only one X was requested and the behavior disappeared completely.

As expected, children's answers tend towards positive assessments with both instruments. To test that the instruments tested the same concepts, we calculated correlations (Pearson's r) between the two formats. Items a and f achieved high correlation (item a: $r(80)=.681$, $p=.000$; item f: $r(82)=.601$, $p=.000$), while the others achieved medium correlation: item b: $r(82)=.482$, $p=.000$, item c: $r(64)=.479$, $p=.000$, item d: $r(65)=.325$, $p=.008$ and item e: $r(65)=.339$, $p=.006$.

BiCo was designed with the intention to mitigate the opinion ceiling effect and therefore to allow for better statistical analysis by providing a lower mean and greater variance, resulting in better differentiation. As Table 3 demonstrates, the mean values were lower for each of the six items. The standard deviation increased for three of the six items, though absolute differences were small. (*n* is smaller for items c, d and e, because those were only administered in the collaborative conditions.)

To test the decrease of means in BiCo compared to Smileyometer statistically, we ran six repeated-measures ANOVAs. Those ANOVAs compared the means of the six concepts that were measured with both item formats over all participants. To control for the multiple comparisons problem, the alpha-level had been set to .008 using a Bonferroni correction. Descriptive results are presented in Table 3.

Table 3.
M (SD) for six items presented in Smileyometer vs. BiCo format.

			Smileyometer		BiCo	
#	Item	n	M	SD	M	SD
(a)	I want to continue	80	.844	.250	.709	.283
(b)	I learned something	82	.720	.253	.642	.275
(c)	I helped my peer	64	.688	.264	.638	.241
(d)	My peer helped me	65	.785	.225	.782	.195
(e)	We collaborated well	65	.885	.183	.803	.211
(f)	I liked the game	82	.762	.257	.732	.255

Table 4.
Percentage of extreme answers with Smileyometer vs. BiCo.

		Smileyometer		BiCo	
#	Item	n	Percentage	n	Percentage
(a)	I want to continue	80	67.5%	82	8.5%
(b)	I learned something	82	30.5%	82	4.9%
(c)	I helped my peer	65	29.2%	65	3.1%
(d)	My peer helped me	66	42.4%	65	12.3%
(e)	We collaborated well	65	67.7%	66	13.6%
(f)	I liked the game	82	43.9%	82	12.2%

Item a revealed a significant main effect of item format with a large effect size: $F(1,79)=31.718$, $p=.000$, $\eta^2=.286$. The same significant effect was found for item b ($F(1,81)=10.251$, $p=.002$, $\eta^2=.112$) and item e ($F(1,64)=8.342$, $p=.005$, $\eta^2=.115$) with medium-to-large effect sizes. Items c, d, and f differed not significantly between item formats. Those findings indicate that BiCo statistically meaningfully lowered the means of three items (a, b, e) compared to Smileyometer. Also in the three other items, the means are lower for BiCo compared to Smileyometer; however, those differences were not statistically significant.

With BiCo, the overall frequency of choosing the most extreme option (i.e. 'very good' in Smileyometer or '1' in BiCo) was lower compared to Smileyometer, as shown by the descriptive data in table 4. Averaged over the six items, the frequency of choosing the most extreme option was 46.9% for Smileyometer, but only 9.1% for BiCo.

Summarized, our findings of significant positive correlations suggest that BiCo and Smileyometer measure the same concepts, pointing to convergent validity between the instruments. Moreover, the means of BiCo were lower for all items compared to Smileyometer; the difference was statistically significant for three out of six items. Finally, for all of the BiCo items children were less likely to choose the most extreme option compared to Smileyometer.

5. Discussion

Surveying children's perception of an (educational) technology is important for developing these technologies further. Research with children as active participants is valuable and possible. To that end, there is a need to develop child-friendly questionnaire methods to appropriately sample children's opinions and experiences. A number of survey instruments exists for children; one particularly popular one is the Smileyometer. Smileyometer already addresses the challenges of surveying children well, but was often associated with opinion ceiling effects in prior studies, minimizing the informative value of the questionnaire results. Furthermore, Smileyometer does not allow for presenting inverted items and it is not well suited for concepts beyond fun. Other survey instruments, like This or That, build on children's ability to draw relative comparisons and show promising results as a tool to adequately sample children's opinions on different facets of a technology experience, avoiding the opinion ceiling effect. However, the comparison instruments usually require a within-subjects research design, which is often not feasible in empirical research. With these prior directions and findings in mind, we developed BiCo to address the opinion ceiling effect of children's technology evaluation, while allowing for an evaluation of only one technology as it is often done in between-subjects studies. This article introduced BiCo, argued for its particular design and provided first data comparing BiCo to an already existing survey instrument for children.

BiCo presents two opposing statements along a continuous bipolar axis, allowing the child participants to draw relative comparisons between them. BiCo can be used in a between-subject research design, where one participant is only exposed to one experimental condition. Having two statements on a bipolar axis allows for switching sides of the expected, typically positive response within a set of items easily. This is a way to reduce the acquiescence bias without having to resort to inverted items, which can be difficult to parse, especially for children. The continuous scale allows for more nuanced responses. Finally, BiCo allows to gather data on subjects outside the area of fun without changing the instrument.

The results of our data comparing BiCo to the well-established Smileyometer instrument suggests that BiCo's theoretical advantages for experimental CCI research show up in practice too: BiCo's and Smileyometer's correlations on the same topics suggest convergent validity between the two instruments for surveying fourth graders. Also our main goal, mitigating the opinion ceiling effect, was met: The means decreased in all six items compared to Smileyometer; for half, this decrease was statistically significant. Furthermore, the frequency of choosing the most extreme answer was lower for BiCo compared to Smileyometer. The results of this initial data indicate BiCo's potential as an instrument to be used in CCI research, however, they come with the limitation that the children used both instruments in the post-questionnaire: While this within-subject design has the advantage of reduced error variance, it could be that exposure to one instrument influenced the response behavior to the other one. In other words, that children had just answered the questions on the same topic with Smileyometer might have influenced how they answered the BiCo questions. However, it has been found that question order effects are less pronounced for younger questionnaire users, because the still developing cognitive skills restrain the ability to incorporate the context of all other items when evaluating the current item (Fuchs, 2005). Considering this, the above-mentioned limitation may be only mildly relevant.

To increase our understanding of BiCo as an instrument in CCI research, future research should be done: BiCo worked well with fourth graders, but fourth graders may be a particularly appropriate audience for BiCo use. They are mentally mature enough to compare two statements with each other, which speaks in favor of BiCo, yet tend to show a high opinion ceiling effect, which is detrimental for Smileyometer. It could well be that younger children are unable to cognitively engage with BiCo and that older children produce better data with Smileyometer. Since age is a critical factor in instrument use (Chambers & Johnston, 2002; Fuchs, 2005), future research should investigate BiCo's utility when used with older or younger children. More generally, BiCo should be applied in differing learning settings and with respondents from various educational, social or cultural backgrounds to assess the generalizability of the results. Also, the mental effort required to use BiCo should be investigated. For instance, how long does it take children to answer a question with BiCo in comparison to Smileyometer? Based on advice on surveying children, we kept the questionnaire short. Children were able to finish it in less than five minutes. For future research, it will be useful to measure at what point children lose interest or find the instrument too taxing. In addition, future research could improve on the design of BiCo. There are findings that labeled scales produce better results than non-labeled scales for children from 10 years on (Borgers, Hox, & Sikkel, 2003). BiCo is scarcely labeled, so one could research if a more exhaustive label layout would improve BiCo. Crafting bipolar statements to contrast is still a largely unexplored topic. How do children cognitively engage with the comparison? Based on that, how do we design good BiCo questions? Creating good questions is a central challenge to surveying children. For instance, in one study (Sim & Horton, 2012), children interpreted ‘Which is a little bit stupid?’ as either being positive or negative. Crafting two statements might have serious advantages: A complex question can be split into two simple statements that are easier for children to parse; short sentences with easy syntax are recommended for the use with children (Bell, 2007). It may also have serious disadvantages: Comparisons could be mentally exhausting and put children off. If BiCo is supposed to be applied in more and broader contexts, we need to know how to translate existing items into the BiCo format while keeping construct equivalence. Based on our findings, this is not trivial. While there is work to be done, this first work suggests that BiCo is a promising instrument for surveying children’s attitudes towards technology in settling the problem of the opinion ceiling effect.

References

Beitchman, J. H. & Corradini, A. (1988). Self-report measures for use with children: a review and comment. *Journal of Clinical Psychology, 44*(4), 477–490. doi:10.1002/1097-4679(198807)44:4<477::aid-jclp2270440402>3.0.co;2-k

Bell, A. (2007). Designing and testing questionnaires for children. *Journal of Research in Nursing, 12*(5), 461–469. doi:10.1177/1744987107079616

Borgers, N., de Leeuw, E., & Hox, J. (2000). Children as respondents in survey research: Cognitive development and response quality 1. *Bulletin de Méthodologie Sociologique, 66*(1), 60–75. doi:10.1177/075910630006600106

Borgers, N., Hox, J., & Sikkel, D. (2003). Response quality in survey research with children and adolescents: The effect of labelled response options and vague quantifiers. *International Journal of Public Opinion Research, 15*(1), 83–94. doi: 10.1093/ijpor/15.1.83

Borgers, N., Sikkel, D., & Hox, J. (2004). Response effects in surveys on children and adolescents: The effect of number of response options, negative wording, and neutral mid-point. *Quality and Quantity, 38*(1), 17–33. doi:10.1023/b:ququ.0000013236.29205.a6

Chambers, C. T. & Johnston, C. (2002). Developmental differences in children’s use of rating scales. *Journal of Pediatric Psychology, 27*(1), 27–36. doi: 10.1093/jpepsy/27.1.27

Chimbo, B., & Gelderblom, H. (2018). TitanTutor: An educational technology solution co-designed by children from different age groups and socio-economic backgrounds. *International Journal of Child-Computer Interaction*, *15*, 13–23. doi:10.1016/j.ijcci.2017.11.004

Colombo, L. & Landoni, M. (2014). A diary study of children's user experience with ebooks using flow theory as framework. *Proceedings of the 2014 Conference on Interaction Design and Children, Aarhus, Denmark*, 135–144. doi: 10.1145/2593968.2593978

De Leeuw, E.D. & Otter, M.E. (1995). The reliability of children's responses to questionnaire items: Question effects in children's questionnaire data. In J.J. Hox, B.F. van der Meulen, J.M.A.M. Janssens, J.J.F. ter Laak, & L.W.C. Tavecchio (Eds.), *Advances in Family Research* (pp. 251-257). Amsterdam: Thesis Publishers.

Fails, J. A., Guha, M. L., & Druin, A. (2012). Methods and techniques for involving children in the design of new technology for children. *Foundations and Trends in Human-Computer Interaction, 6*(2), 85–166. doi:10.1561/1100000018

Fowler, A. (2013). Measuring learning and fun in video games for young children: A proposed method. *Proceedings of the 12th International Conference on Interaction Design and Children, New York, New York, USA*, 639-642. doi:10.1145/2485760.2485879

Fraser, S., Flewitt, R., & Hammersley, M. (2014). What is Research with Children and Young People? In A. Clark, R. Flewitt, M. Hammersley, & M. Robb (Eds.), *Understanding Research with Children and Young People* (pp. 34–50). SAGE. doi:10.4135/9781526435637.n3

Fuchs, M. (2005). Children and Adolescents as Respondents. Experiments on Question Order, Response Order, Scale Effects and the Effect of Numeric Values Associated with Response Options. *Journal of Official Statistics*, *21*(4), 701–725.

Hall, L., Hume, C., & Tazzyman, S. (2016). Five degrees of happiness: Effective smiley face Likert scales for evaluating with children. *Proceedings of the 15th International Conference on Interaction Design and Children, Manchester, United Kingdom*, 311–321. doi:10.1145/2930674.2930719

Hanna, L., Neapolitan, D., & Risden, K. (2004). Evaluating computer game concepts with children. *Proceedings of the 2004 Conference on Interaction Design and Children: Building a Community, Maryland*, 49–56. doi:10.1145/1017833.1017840

Harris, A., Rick, J., Bonnett, V., Yuill, N., Fleck, R., Marshall, P., & Rogers, Y. (2009). Around the table: Are multiple-touch surfaces better than single-touch for children's collaborative interactions? In C. O'Malley, D. Suthers, P. Reimann, & A. Dimitracopoulou (Eds.), *Computer Supported Collaborative Learning Practices - CSCL2009 Conference Proceedings* (pp. 335–344). Rhodes, Greece: International Society of the Learning Sciences.

Horton, M. (2013). *Improving Validity and Reliability in Children's Self Reports of Technology Use* (Doctoral dissertation, University of Central Lancashire). Retrieved from http://clok.uclan.ac.uk/9234/

Horton, M., Read, J. C., & Sim, G. (2011). Making your mind up?: The reliability of children's survey responses. *Proceedings of the 25th BCS Conference on Human-Computer Interaction, Newcastle-upon-Tyne, United Kingdom*, 437–438.

Johnson, R., Shum, V., Rogers, Y., & Marquardt, N. (2016). Make or shake: An empirical study of the value of making in learning about computing technology. *Proceedings of the 15th International Conference on Interaction Design and Children, Manchester, United Kingdom*, 440–451. doi:10.1145/2930674.2930691

Kano, A., Horton, M., & Read, J. C. (2010). Thumbs-up scale and frequency of use scale for use in self reporting of children's computer experience. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries, Reykjavik, Iceland*, 699–702. doi:10.1145/1868914.1869008

Law, E. L.-C., Watkins, D. E., Barwick, J. P. L., & Kirk, E. S. (2013). An experiential approach to the design and evaluation of a gamified research tool for law in children's lives. *Proceedings of the 15th International Conference on Interaction Design and Children, Manchester, United Kingdom*, 322–333. doi:10.1145/2930674.2930722

Looije, R., Neerincx, M. A., & de Lange, V. (2008). Children's responses and opinion on three bots that motivate, educate and play. *Journal of Physical Agents, 2*(2), 13–20. doi:10.14198/jopha.2008.2.2.03

Marsh, H. W. (1986). Negative item bias in ratings scales for preadolescent children: A cognitive-developmental phenomenon. *Developmental Psychology, 22*(1), 37–49. doi:10.1037//0012-1649.22.1.37

Moosbrugger, H. & Kelava, A. (2008). *Testtheorie und Fragebogenkonstruktion*. Berlin: Springer. doi:10.1007/978-3-540-71635-8

Nielsen, J. (1993). *Usability Engineering*. San Francisco: Morgan Kaufmann Publishers Inc.

Read, J. C. (2008). Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. *Cognition, Technology & Work, 10*, 119–128. doi:10.1007/s10111-007-0069-9

Read, J. C. & MacFarlane, S. (2006). Using the Fun Toolkit and other survey methods to gather opinions in child computer interaction. *Proceedings of the 2006 Conference on Interaction Design and Children, Tampere, Finland*, 81–88. doi:10.1145/1139073.1139096

Read, J. & Markopoulos, P. (2013). Child–computer interaction. *International Journal of Child-Computer Interaction, 1*(1), 2–6. doi:10.1016/j.ijcci.2012.09.001

Rick, J. (2012). Proportion: a tablet app for collaborative learning. *Proceedings of the 11th International Conference on Interaction Design and Children, Bremen, Germany*, 316–319. doi:10.1145/2307096.2307155

Schmitt, L. J., & Weinberger, A. (2019). Fourth graders' dyadic learning on multi-touch interfaces—versatile effects of verbalization prompts. *Educational Technology Research and Development*, *67*(3), 519–539. doi:10.1007/s11423-018-9619-5

Scott, J. (2000). Children as Respondents: The Challenge for Qualitative Researchers. In P. Christensen, & A. James (Eds.), *Research with Children: Perspectives and Practices* (pp. 98–119). London: Falmer Press.

Scott, J., Brynin, M., & Smith, R. (1995). Interviewing children in the British household panel survey. In J.J. Hox, B.F. van der Meulen, J.M.A.M. Janssens, J.J.F. ter Laak, & L.W.C. Tavecchio (Eds.), *Advances in Family Research* (pp. 259-266). Amsterdam: Thesis Publishers.

Sim, G. & Cassidy, B. (2013). Investigating the fidelity effect when evaluating game prototypes with children. *Proceedings of the 27th International BCS Human Computer Interaction Conference, London, United Kingdom*, 53:1-53:6.

Sim, G. & Horton, M. (2012). Investigating children's opinions of games: Fun toolkit vs. this or that. In Proceedings of IDC '12 (pp. 70–77). New York: ACM. doi:10.1145/2307096.2307105

Sim, G., Horton, M., & McKnight, L. (2016). iPad vs paper prototypes: Does form factor affect children's ratings of a game concept? *Proceedings of the 11th International Conference on Interaction Design and Children, Bremen, Germany*, 190–195. doi:10.1145/2930674.2930720

Sim, G., MacFarlane, S., & Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. *Computers & Education, 46*(3), 235–248. doi:10.1016/j.compedu.2005.11.021

Smith, K., & Platt, L. (2013). *How do children answer questions about frequencies and quantities? Evidence from a large-scale field test*. London: Centre for Longitudinal Studies.

Zaman, B., Abeele, V. V., & De Grooff, D. D. (2013). Measuring product liking in preschool children: An evaluation of the Smileyometer and This or That methods. *International Journal of Child-Computer Interaction, 1*(2), 61–70. doi:10.1016/j.ijcci.2012.12.001

Study III:

Schmitt, L. J., & Weinberger, A. (2019). Fourth graders' dyadic learning on multi-touch interfaces – versatile effects of verbalization prompts. *Educational Technology Research and Development, 67*(3), 519-539. doi:10.1007/s11423-018-9619-5

This is a post-peer-review, pre-copyedit version of an article published in Educational Technology Research and Development. The final authenticated version is available online at: https://doi.org/10.1007/s11423-018-9619-5.

Fourth graders' dyadic learning on multitouch interfaces—versatile effects of verbalization prompts

Lara Johanna Schmitt & Armin Weinberger

Abstract

Multi-touch interfaces allow for direct and simultaneous input by several co-present learners and afford hands-on learning experiences. Additional scaffolding for strategic behavior and/or verbalizations may constructively complement collaborative learning with a multi-touch device. In this study, the tablet app ''Proportion'' is supposed to enable two novices (about 10 years old) to collaboratively construct an understanding of proportional relations. In a 2 × 2 factorial design (n = 162), effects of enriching Proportion with strategy prompts (with/without) and verbalization prompts (with/without) on multi-modal processes as well as near and far transfer learning gains have been investigated. The process variables include task focus, positive and negative emotions, and quality of dialogue (transactivity, epistemic quality). We found a general improvement in near transfer task types over all conditions without the two prompt types further affecting learning gains. While the strategy prompts did not significantly affect processes or outcomes, the verbalization prompts had versatile effects on learning processes: On one hand, quality of talk was improved, on the other hand, task focus and emotions were negatively affected.

Keywords: Collaborative learning, Embodiment, Proportional reasoning, Scaffolding, Tablets

Multi-touch interfaces for collaborative embodied learning

Multi-touch interfaces on smartphones and tablets play a major role in how we consume media today, and interacting with multi-touch devices is enjoyable for most learners (e.g., Alvarez, Brown, & Nussbaum, 2011). They allow for co-present collaborative learning and, specifically, afford equal, simultaneous, and direct manipulation of digital objects in a learning environment (Alvarez et al., 2011; Mercier, Higgins, & da Costa, 2014; Roschelle et al., 2010). Direct manipulation of digital objects is supposed to reduce the "cognitive distance between intent and execution" (Rick, 2012, p. 316) and allows for "expanding the scope of what can be experienced" (Ottmar & Landy, 2017, p. 72). From an embodied cognition point of view, experiences with the environment form multi-modal mental representations (Barsalou, 2008). Thus, the body is an additional resource that extends and complements cognitions (Abrahamson, 2017; Davidson & Ryberg, 2017). Multi-touch learning environments can afford "embodied learning experiences" to different degrees in the sense that they involve bodily movement and hands-on activities with representations of knowledge (Ottmar & Landy, 2017; Sakr, Jewitt, & Price, 2014; Schneps et al., 2014). Embodied learning experiences, e.g. physically moving algebraic representations on a tablet, have been shown to positively affect discussion (Ottmar & Landy, 2017), retention of knowledge (Cook, Mitchell, & Goldin-Meadow, 2008) and reasoning (Reinholz, Trninic, Howison, & Abrahamson, 2010). However, these hands-on learning activities may also provoke off-task behavior and distract from the actual learning goals (e.g., Danish, Enyedy, Saleh, Lee, & Andrade, 2015), thus come at the cost of verbalization, abstraction, and reflection of knowledge.

In order to work towards a profound understanding of the mechanisms of collaborative learning with multi-touch devices, we consider both cognitive and emotional processes of learning represented through bodily and verbal data sources, taking into account the underlying assumptions of how these processes relate to learning. Positive emotions, e.g. enjoyment or pride, correlate positively and negative emotions, e.g. boredom or hopelessness, correlate negatively with learning (Pekrun, Goetz, Titz, & Perry, 2002). Positive emotions also sustain children's motivation to interact with technology (Sim, Cassidy, & Read, 2013). On the cognitive side, task focus is an indicator for student engagement with the learning environment (Lehman, Matthews, D'Mello, & Person, 2008), as well as a prerequisite and predictor for learning (Baker and Lund, 1997; Cohen & Lotan, 1995). Beyond task focus, two qualities of learners' talk, transactivity and epistemic quality, have been identified as particularly conducive to learning (Weinberger & Fischer, 2006). Transactivity regards the extent to which participants are referring to their partner's contributions; epistemic quality regards how learners use and connect the relevant concepts. In our analyses, we will build up on the framework by Weinberger and Fischer (2006) to analyze these qualities. This framework identified qualitatively different transactive and epistemic moves. For example, purely elaborating on a subject is less transactive than conflict-oriented consensus building, during which learners integrate and challenge their learning partner's contribution. Regarding epistemic quality, the framework distinguishes non-epistemic activities, e.g., to coordinate the environment or off-topic talk, from epistemic activities that focus on the problem space or the conceptual space and the relations between them. These cognitive process variables can be assessed through both, learners' verbalizations indicating how learners understand and apply concepts, as well as bodily indicators for learners' focus on the task (Deater-Deckard, El Mallah, Chang, Evans, & Norton, 2014; Sakr, Jewitt, & Price, 2016). To analyze verbalizations, we will adopt and develop the framework by Weinberger and Fischer (2006) to suit our purpose of investigating children's collaborative learning with a multi-touch device.

Scaffolding CSCL

Groups of learners often do not develop their full potential and commonly display process losses (Kerr & Tindale, 2004). Furberg (2016) identified *procedural* and *conceptual* challenges that students face while engaging in CSCL (computer supported collaborative learning). Procedural challenges refer to difficulties in carrying out the concrete learning tasks at hand, whereas conceptual challenges refer to difficulties in connecting the concrete learning tasks to more abstract concepts of the learning domain (Furberg, 2016). These challenges afford specific types of additional scaffolding for collaborative learners (e.g., Borge & White, 2016). While the role of the teacher as a mediator and facilitator of collaborative learning in the classroom has been rediscovered and emphasized (Caballero et al., 2014), the support can also be embedded in the software. The following paragraphs will elaborate on how the aforementioned two types of challenges in CSCL can be addressed by two types of scaffolding, strategy scaffolding addressing procedural, and verbalization scaffolding addressing conceptual challenges.

Strategy scaffolding is designed to help learners in proceeding in the task solution processes by guiding them to better understand what they have to do, what is a sensible sequence of activities, or on what to focus (Jackson, Krajcik, & Soloway, 1998; Sharples et al., 2015). Prompting learners to invest time into thinking and planning activities for understanding the problem instead of orientation towards minimal requirements positively affects learning processes and outcomes, for example expert-like problem solving (Schoenfeld, 1987, 1992) and mathematical modelling competency (Schukajlow, Kolter, & Blum, 2015). Structuring interactions increases task-focus (Baker & Lund, 1997). Also, it has been argued that supporting learners by facilitating the problem solving processes may foster confidence of success and therefore motivation (Belland, Kim, & Hannafin, 2013). *Strategy scaffolding* can be realized by providing additional structures, for example by breaking down a complex task into manageable steps (Schukajlow et al., 2015; Sharples et al., 2015), or by providing learners with pre-set "communication acts" to choose from (e.g., "Where do we start?"; Baker & Lund, 1997, p. 183). *Strategy scaffolding* can also be realized by prompting for strategic behavior ("You really should think about making a plan before starting to build your model", Jackson et al., 1998, p. 189).

Verbalization scaffolding aims to facilitate elaboration of the learning material (King, 1990). To learn successfully, CSCL participants need to actively engage in building and sharing their knowledge (Baker & Lund, 1997; Roschelle & Teasley, 1995), but often stay on the surface (e.g., Falloon & Khoo, 2014), or even wander off completely (e.g., Danish et al., 2015). Reiser (2004) argued that "problematizing aspects of subject matter" (p. 287) can be realized by requesting additional explanations. Problematizing is assumed to help students focusing on the task and to support effective learning (Reiser, 2004). Empirical findings are that prompting learners' verbalizations increased elaboration of the learning material and learners' understanding (King, 1990). Examples for verbalization prompts are "Explain why…", "How are…and…similar?" (King, 1990, p. 669), or "What does all this mean to you?" (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001, pp. 527-528). Children aged 6-7 years had a more interactive and elaborate discussion when prompted by "review" and "thinking" question prompts (Gelmini-Hornsby, Ainsworth, & O'Malley, 2011, p. 582), hence displaying higher levels of transactivity and epistemic quality. Joint verbalization, i.e. building and sharing arguments together, may productively coalesce with embodied learning experiences and thus, further translate between embodied experience and abstract conceptualization. However, it has been argued that the sudden appearance of reflection prompts may disturb the flow and thus decrease motivation and learning outcomes in game-based learning environments (Wouters et al., 2015). This argument was partially confirmed by Chen and Law (2016) who found that additional question prompts in a tablet-based learning environment enhanced seventh graders' learning outcomes, but decreased their motivation.

So potentially, there is a need for scaffolding learning on multi-touch interfaces for *strategic* behavior, *verbalization* of underlying concepts, or both. To our best knowledge, there are no prior studies that experimentally investigate these two prompt types in combination in a multi-touch learning environment. Would additional scaffolding constructively complement or rather disrupt embodied learning experiences (see also, Schneps et al., 2014)? While positive effects, i.e. improved cognition and discussion quality, can be assumed, prompts may also put learners off.

Proportional Reasoning

Proportional reasoning is a central topic in mathematics (Boyer, Levine, & Huttenlocher, 2008) and defined as "reasoning with ratios, rates, and percentages" (Jitendra, Star, Rodriguez, Lindell, & Someki, 2011, p. 731), or "understanding the multiplicative relationships between rational quantities" (Boyer et al., 2008, p. 1478). Children typically experience difficulties in proportional reasoning, particularly with fractions (Mix, Levine, & Huttenlocher, 1999). These difficulties have been explained by a lack of senso-motorical experiences with proportions (Reinholz et al., 2010). Typical misconceptions that children face in proportional reasoning include the application of counting strategies (Gelman, Cohen, & Hartnett, 1989), especially when concrete units are presented (Boyer et al., 2008), as well as addition rules (Mix et al., 1999) to proportions. Furthermore, it is hard for children to form correct proportional representations from discrete units, i.e. to build a relative relationship between numerator and denominator of a fraction (Boyer et al., 2008; Mix et al., 1999).

Fostering proportional reasoning requires effective learning set-ups, e.g. building on collaborative hypothesis testing (Ellis, Klahr, & Siegler, 1993). Martin et al. (2015) emphasize the positive role of "opportunities to engage in learning environments that repeatedly demonstrate and allow practice with these concepts [i.e., fractions] without explicit instruction." (p. 629). Recent developments also aim at including embodied learning experiences. Moving the hands up and down in the air to represent proportional relations showed potential to support students', grade four to six, pre-symbolic reasoning about proportions (Mathematical Image Trainer; Reinholz et al., 2010). In this vein, we have developed an iPad app called "Proportion" that aims at improving children's proportional reasoning: Rather than dealing with abstract numbers, learners can directly manipulate volumes associated with the numbers' symbolic representations. The numbers and their proportional relationships become visible and “graspable”.

Research questions and hypotheses

The present study contributes to the ongoing debate about collaborative learning in multi-touch learning environments. Scaffolding learning on multi-touch interfaces could be both: counterproductive, thus ruining and disrupting the bodily, hands-on learning experience, or complementary, thus improving and supplementing it with processes of encoding that are necessary for memorization and learning. Our main aim is to investigate the differential effects of scaffolding for strategic behavior and/or reflective dialogue on learning processes and outcomes in a multi-touch environment affording embodied learning experiences. We deploy and investigate the effects and the interaction of strategy prompts / “STRAT” and verbalization prompts / “VERB”. The two prompt types have been varied in a 2×2 design with repeated measures.

- **Hypothesis 1:** STRAT prompts and VERB prompts and their combination will result in higher task focus. Structuring the task with the STRAT prompts should help learners to make progress and stay on track. Requesting verbalization of underlying concepts with the VERB prompts should direct learners’ attention to relevant task features, thus also foster task focus.
- **Hypothesis 2:** The STRAT prompts will develop positive emotions; the VERB prompts will rear negative emotions. Providing strategies (STRAT) is supposed to advance positive emotions like enjoyment, because task progression is facilitated. Making learners verbalize (VERB) is supposed to result in negative emotions like frustration or anger, because task progression is slowed down by difficult verbal tasks.
- **Hypothesis 3:** The VERB prompts will enhance the quality of dialogic interactions. Explicitly requesting learners to externalize their knowledge and engage in discussions should result in higher transactivity and higher epistemic quality.
- **Hypothesis 4:** STRAT prompts and VERB prompts and their combination will result in higher learning gains, regarding both near and far transfer task types. The internalization of task solving strategies suggested by the STRAT prompts should have a positive impact on learning. Higher-order verbalizations as scaffolded by the

VERB prompts should impact learning positively by promoting deep elaboration and abstract thinking.

Methods

Sample

Participants were fourth graders (50% male) with a mean age of 10 years and 4 months (*SD* = 6.7 months) from seven primary schools in Germany. All participants had a consent form signed by their legal representatives, informing on data collection and analyses. In total, n = 162 participants took part in the experiments, being tested in four different experimental conditions: control, STRAT, VERB, and VERB-STRAT. To arrive at a sample representing schools of differently favorable backgrounds, school districts were classified beforehand according to the percentage of pupils who had been granted access to Gymnasium, the highest secondary school level in Germany, which is eventually granting access to University education (*1. Bildungsbericht*, 2012): Below average (13-36% continuing at high-school), average (50-55% continuing at high-school), and above average (66-68% continuing at high-school). χ^2-tests and ANOVAs did not reveal statistically significant differences between conditions regarding the control variables gender and gender-dyad composition (boy-boy, girl-girl, boy-girl), school background, handedness, experience with or owning of a multi-touch device, pre-test and dyads' homogeneity (coefficient of variation; Weinberger, Stegmann, & Fischer, 2007) in the pre-test.

Due to technical difficulties, there was some data loss regarding the analysis of the learning *processes*: One dyad's video (control condition) could not be analyzed at all due to the video's bad quality. Furthermore, six dyads' videos (all participants from the VERB condition) could only be partially analyzed: In four cases, the app crashed during the experiment after 25, 25, 11, and 30 minutes respectively. In case of app crashes, the experimenter restarted the app and quickly leveled up to the problem the participants were previously working on before returning the tablet to the participants. Though we included non-verbal behavior (task focus, positive and negative emotions) until the crash, we did not quantify the quality of the dialogue. In two other cases, the audio quality was bad and we analyzed non-verbal behavior, but not the dialogues. To control for systematic bias caused by the data loss, we tested statistically if the reduced sample size impacts the comparability between groups regarding the above mentioned control variables. We found the groups to be still comparable in all control variables, except for an inequality of distribution in the variable "handedness" (Fisher's exact test, p = .014). However, both the original sample as well as the reduced sample show the same pattern of inequality (three left-handed participants in each, VERB and control, and zero left-handed people in each, STRAT and VERB-STRAT), which we think is a negligible floor effect.

Material

The learning environment Proportion is an iPad application (Rick, 2012). The interface is designed to afford incorporation of hand / arm movements, i.e. aiming at actively experiencing proportional relations through direct manipulation. Proportion challenges learners to solve increasingly difficult levels and provides positive feedback on success, overall promoting a game-like user experience (Rick, Kopp, Schmitt, & Weinberger, 2015). The app is subdivided into 21 levels. Each level comprises between 5 to 23 problems. In all experimental conditions, learners can control a blue and an orange bar, which are positioned vertically next to each other. Each bar is assigned a number. The bars need to be in the right relation indicated by the numbers in order to solve the problem. This can be achieved by resizing the bars.

Figure 1. The interface of Proportion (a), example of displaying one STRAT (b) and one VERB (c) prompt, two children collaboratively using Proportion (d)

See figure 1 (a) for an example: In this case, the bars need to be adjusted so that the orange bar's height would be 2/3 compared to the blue one. Once a problem is solved, the numbers of the next problem automatically appear. The levels are characterized by different task types, e.g. integers, fractions with same vs. different denominator. A small owl acts as a pedagogical agent and provides feedback, e.g. announces "correct" once a problem is solved.

This pedagogical agent also displays the varied STRAT and VERB scaffolding, see figure 1 (b) and (c); each prompt type has three versions (A, B, or C); for an overview on the prompts, see table 1.

Table 1
Overview on prompts

Prompt type	Level	Prompt text	Theoretical foundation
STRAT	1, 4, 7, 10, 13, 16, 19	A: Tip for all tasks: What is higher, orange or blue? First say it out loud, then you do!	Guiding learners by providing additional structure and fostering strategic behavior that can applied throughout the learning phase (Jackson et al., 1998; Schoenfeld, 1987;
	2, 5, 8, 11, 14, 17, 20	B: Tip for all tasks: First think and provide an estimate, then set the bars' correct height!	
	3, 6, 9, 12, 15, 18, 21	C: Tip for all tasks: If the task is hard and you're stuck, what might help is to discuss and talk!	Schukajlow et al., 2015; Sharples et al., 2015)
VERB	1, 4, 7, 10, 13, 16, 19	A: Explain to your learning partner: What did one need to do in order to solve the task?	Supporting knowledge construction by eliciting high-quality verbalizations, requesting concrete
	2, 5, 8, 11, 14, 17, 20	B: Describe to your learning partner: What could one learn in this task?	explanations to pre-determined questions (Chi et al., 2001; King, 1990; Reiser, 2004)
	3, 6, 9, 12, 15, 18, 21	C: Explain to your learning partner: What do all of these tasks have in common?	

We administered two *questionnaires*. The first questionnaire, prior to the learning period, consisted of 9 items and collected socio-demographic data, previous experiences with multi-touch devices, and attitudes towards math, school in general, and learning collaboratively vs. individually. The second questionnaire, after the learning period, consisted of 14 items measuring participants' acceptance of the app, subjective learning gain, and aspects of the collaboration.

The math tests consisted of items related to fractions and proportions; the items were classified as requiring lower vs. higher levels of transfer (near transfer tasks vs. far transfer tasks). The near transfer tasks (10 sub-tasks / points) were designed to capture the strategies that were used to progress within Proportion. The far transfer tasks (18 sub-tasks / points) aimed at capturing knowledge on proportions and fractions more broadly. Because no standardized tests on this domain and age group were available, the far transfer tasks were adapted from another research project on proportional thinking (Rick, Rogers, Haig, & Yuill, 2009) and subsequently adapted to suit our subjects. An item analysis on the pre-test led us to exclude 7 items (2 from near transfer and 5 from far transfer tasks) because of extremely low or high item difficulty and/or insufficient discriminatory power. Cronbach's Alpha in the revised version of the math-test reached .76 for near transfer tasks (now 8 sub-tasks / points) and .68 for far transfer tasks (now 13 sub-tasks / points).

Seven iPads of the second generation were used for the experiments. Video cameras and microphones recorded participants' interactions with each other and Proportion.

Experimental Procedure

The experimental procedure followed a 2×2 design with repeated measures (pre-test – post-test) and has been carried out by one of several trained experimenters. Experiments took place inside the respective schools, replacing two regular school lessons (90 minutes). We scheduled the experiments together with the teachers beforehand without the teachers being present during the experiments to keep any external influence to a minimum. Students' performance within the study was not affecting their regular math grades. Participants were randomly assigned to conditions by lot. After a general welcoming and introduction to the learning session, the participants individually filled in the first questionnaire and the pre-math test. Next, the students worked collaboratively with the Proportion app for 40 minutes, see figure 1 (d); this phase was video-taped. After a 5 minutes break, the participants individually filled in the second questionnaire and post- math test. The completion of questionnaires was not time-limited. The experiment continued after everyone was done, usually after around three to five minutes. The completion of the pre- and the post- math test was time-constrained to 10 minutes of time for each test. Mostly, the participants were done with the test well before the 10 minutes ran out.

Variables and data analysis

The outcome variables learning gain in near and far transfer tasks refer to the changes from the pre- to the post- math test.

We analyzed the videos of the participants learning with Proportion both qualitatively and quantitatively. We coded the following process variables: emotions, task focus, and quality of talk. To capture these variables, we created and refined coding schemes in several iterations. The coding schemes are based on previous work, where the amount of participation, epistemic quality and transactive moves (Weinberger & Fischer, 2006), as well as facial expressions, speech and gestures as indicators for emotions (e.g., Lehman et al., 2008; Sim, MacFarlane, & Read, 2006) have been coded. The dependent variables have been aggregated, i.e. averaged, on the dyad level as we cannot assume statistical independence of the dyadic learners.

In total, we surveyed 162 participants; the result was 81 videos of 40 minutes each, thus 54 hours of video material. Regarding the analysis of the process variables, we chose to sample every second level (levels 1, 3, 5, …, 21), and within each level we analyzed every second problem. This allowed us to focus on continuous interactions during typical problems throughout the learning phase. The samples covered reactions to successful problem solving, reactions to the prompts for the prompted conditions, and the problem solving process itself. The same problems have been analyzed for all dyads and the video samples were transcribed by trained assistants. Training for achieving inter-rater reliability was done with different videos than the ones that were used to actually check the inter-rater reliability. It was also ensured that the distribution of coding assistants on videos from different conditions was approximately equal.

In terms of non-verbal behavior, the variables positive emotions, negative emotions and task focus have been analyzed. The focus was on gestures for the emotions and eye contact for task focus. Well understandable utterances were taken into account in case the gestures were not clear-cut. In table 2, the criteria for coding of the emotion variables are described. Also, other gestures and actions that were explicitly not considered to be indicators (too ambiguous) are listed. Only emotions during on-task phases have been taken into account. For securing reliable analysis, we opted for a simple distinction between positive and negative emotions (e.g., Frijda, 1988; Linnenbrink & Pintrich, 2002). According to Frijda (1988), positive vs. negative emotions result from events that have the potential to further vs. threaten one’s goal achievement.

Table 2

Overview on coding criteria for emotions

Variable	Indicators	Not counted
Positive emotions	- clapping into one's own or the learning partner's hands - throwing hands up in the air - clenching the fist - showing thumbs-up	- change of body position - soft / indifferent hand movements - scratching the forehead - folding arms - shaking the head - knocking on the iPad or the table - letting the hand(s) fall down - showing the index finger - smiling / laughing
Negative emotions	- threatening the iPad - facing the palms upwards - dismissive hand gesture - face-palming	

Regarding task focus, we drew on bodily indicators for learners' focus on the task (Deater-Deckard et al., 2014; Sakr et al., 2016), see table 3, and coded instances of off-task behavior, i.e. neither active or passive engagement with the problem. As can be seen in table 3, participants would sometimes interact with participants outside their own group. Though we instructed the participants before the experiments to stick to their assigned groups, we could not entirely prevent them of occasionally communicating their level progress to other classmates; this behavior was observed in approximately 1/3 of the experimental sessions; incidents were brief (< 5 sec).

Table 3

Overview on coding criteria for task focus

Variable	On-task behavior	Off-task behavior
Task Focus	- active manipulation of the bars in the app - monitoring of the learning partner manipulating the bars while the eyes are focused on the iPad - verbally interacting with the learning partner - interactions with the experimenters (asking questions related to Proportion, or listening to and watching the experimenters in case they gave instructions)	- looking around the classroom - looking into the camera - interactions with participants outside the own group

Per analyzed problem, we coded how many times each non-verbal behavior (positive emotions, negative emotions, off-task behavior) occurred, i.e. we observed one of the indicators (table 2 and table 3). Coding of non-verbal behavior has been performed by two trained coders. Inter-rater reliability has been assessed for 10% (*n* = 16 participants) of the analyzed sample using Krippendorff's α (Krippendorff, 2012). The α-values reached α = .81 for positive emotions, α = .85 for negative emotions and α = .92 for task focus. In total, we coded 1311 instances of non-verbal behavior (*n* = 160 participants). For the statistical analysis, we compared the average number of off-task events, positive and negative emotions per problem.

In terms of quality of talk, the variables transactivity and epistemic quality have been examined. This coding scheme is largely based on the framework to analyze argumentative knowledge construction in computer-supported collaborative learning (Weinberger & Fischer, 2006). We measured transactivity as limited to the interactions between the learning partners, interactions with the pedagogical agent in the system were not taken into account, and also included verbal reactions to non-verbal behavior / actions with the app. Table 4 lists the categories for transactivity.

Table 4
Categories for transactivity

Category	Description of category	Examples
Externalization (Ext)	Knowledge / thoughts are being externalized without reference to the learning partner	“There are no numbers anymore, it is always the same.”
Externalization as reaction (ExtR)	Reaction to an elicitation without it being an acceptance, refusal, integration or conflict-oriented consensus building	A; “Where?” B: “Here, at the line”
Acceptance (Acc)	Simple and short acceptance of partners’ statement	“Yes”, “OK”, “Good”
Refusal (Ref)	Simple and short rejection of partners’ statement or action	“No”, “Stop”, “Wait”
Elicitation (Eli)	Asking a question or provoking a (verbal or activity-oriented) reaction	“The 3 needs to be higher, doesn’t it?”
Integration (Int)	Elaborated acceptance incl. repetition or rephrasing of what’s been said by the partner (no new content)	A: “Always half of it” B: “Always half, yes”
Conflict-oriented consensus building (COC)	Reference and modification, extension, relativization or counter argumentation to what has been said by the partner	A: “This needs to go higher” B: “You have to go lower. It needs to be three times the size”

We measured epistemic quality in terms of the content of participants’ utterances: Are the utterances off- or on-topic and are they a pure regulation of their interaction or (different levels) of actual task-related explanations? Table 5 lists the categories for epistemic quality.

Table 5
Categories for epistemic quality

Category	Description of category	Examples
Off-topic utterance (OT)	Utterances not dealing with the iPad or the app	"I am hungry"
On-topic: regulation of the interaction (RegIA)	Dialogue, technology, task or learning partner are regulated; structuring of conversation without deeper meaning	"well…", "Give me the iPad", "Correct", reading aloud a prompt
On-topic: concrete task-related regulation (TaskIA)	Hints or explanations that are very close to what one can see and perceive	"I need to stop here", "this is 3 and this is 1"
On-topic: abstract content-related regulation (ContIA)	Hints or explanations that refer to abstract knowledge, more than what one can see or perceive	"double it", "this is a third"
On-topic: procedural knowledge / strategies (also non-verbal) (Proc)	Strategies like counting, measuring, applying the prompts of the pedagogical agent	"1 2 3 4…", "First estimate the height, then set the bars!"

Coding of the dialogues has been performed by four trained coders. Inter-rater reliability has been assessed for 7% (n = 10 participants) of the analyzed sample using Krippendorff's α (Krippendorff, 2012). The α-values reached α = .77 for transactivity and α = .88 for epistemic quality. In total, we coded 3564 dialogue turns (n = 148 participants). The categories of transactivity and epistemic quality had been conceptualized as ordinally scaled in prior work (e.g., Berkowitz & Gibbs, 1983; Teasley, 1997; Weinberger & Fischer, 2006). While in previous work individual categories were selected for the statistical analysis (e.g., Berkowitz & Gibbs, 1983), we decided to build global scores taking into account all categories. To do so, the categories were aggregated and weighted according to their presumed relative quality. Differently weighing intervals between categories aims to correspond to the qualitatively different conceptualization of the respective transactive and epistemic moves (Berkowitz &

Gibbs, 1983; Teasley, 1997; Weinberger & Fischer, 2006): Transactivity score = 1*acceptance + 1*rejection + 2*elicitation + 4*integration + 8*conflict-oriented consensus building. Epistemic quality score = 1*concrete task-related regulation + 4*abstract content-related regulation + 8*strategies/procedural knowledge. Externalization has not been counted as it is not transactive. Off-topic utterances and regulation of the interaction have not been counted because they lack epistemic content.

Results

Hypothesis 1: Effects on task focus

In total, we observed 805 instances of off-task behavior throughout the entirety of the analyzed sample, which corresponds to on average 5.03 (SD = 3.93) instances of off-task behavior per participant. Hypothesis 1 stated a main effect for both STRAT and VERB as well as an interaction between the prompts on task focus. A two-factorial ANOVA revealed no statistically significant main effect of STRAT (F (1,76) = .159, p = .691), but a highly significant main effect of VERB: F (1,76) = 18.190, p = .000, η^2 = .19. However, this highly significant effect is contrary to our hypothesis, as off-task behavior is actually reinforced, and not reduced, with the presence of the VERB prompt, see figure 2. The interaction of the prompts was not statistically significant: F (1,76) = 2.451, p = .122.

Hypothesis 2: Effects on emotions

In total, we observed 274 positive emotions and 232 negative emotions throughout the entirety of the analyzed sample, which corresponds to on average 1.71 (SD = 1.29) positive emotions and 1.45 (SD = 1.65) negative emotions per participant. Hypothesis 2 stated more positive emotions for STRAT and more negative emotions for VERB. A two-factorial ANOVA revealed no statistically significant effect of STRAT on positive emotions: F (1,76) = 1.919, p = .170. The effect of VERB on negative emotions is statistically significant with F (1,76) = 7.019, p = .010, η^2 = .09, see figure 2.

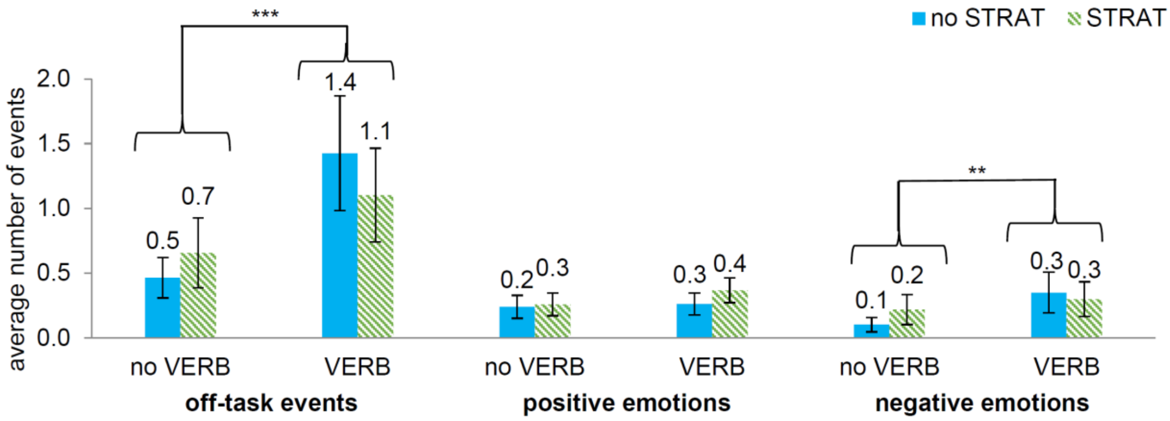


Figure 2. Average number of off-task events, and positive and negative emotions per problem analyzed; main effect of VERB on off-task behavior and negative emotions. Error bars represent 95% CIs. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$.

Hypothesis 3: Effects on quality of dialogue

On average, the participants produced 24.08 (*SD* = 12.26) dialogue turns throughout the entirety of the analyzed sample. In general, off-topic talk was very rare. Most of the participants' epistemic activities were focusing on regulating their task progress. Regarding transactivity, we mostly observed externalizations, but also higher-level verbalizations like elicitation or conflict-oriented consensus building. Descriptive data for the total numbers, means (*SD*) and frequencies of each category of epistemic quality and transactivity are presented in table 6.

Table 6
Observed total numbers, means (SD) per participant, and frequency percentages for epistemic quality and transactivity categories

Epistemic quality				Transactivity			
Category	Total number	M (*SD*)	Frequency percentage	Category	Total number	M (*SD*)	Frequency percentage
Off-topic utterance	43	.3 (.7)	1.4 %	Externalization	1494	10.1 (5.1)	42.8 %
Regulation of the interaction	2186	14.8 (8.4)	59.8 %	Externalization as reaction	86	.6 (.8)	2.5 %
Concrete task-related regulation	909	6.1 (3.5)	25.9 %	Acceptance	451	3.0 (2.7)	11.3 %
Abstract content-related regulation	168	1.1 (1.2)	4.3 %	Refusal	277	1.9 (1.5)	8.2 %
Procedural knowledge / strategies	99	.7 (.8)	3.2 %	Elicitation	715	4.8 (2.9)	20.0 %
				Integration	54	.4 (.6)	1.3 %
				Conflict-oriented consensus building	328	2.2 (2.2)	8.6 %
Not codable	159	1.1 (1.0)	5.4%	Not codable	159	1.1 (1.0)	5.4%

Hypothesis 3 stated a main effect of VERB on the quality of dialogic interactions, regarding transactivity as well as epistemic quality. Two-factorial ANOVAs were conducted. Regarding transactivity, there was a significant main effect of VERB (F (1,70) = 7.241, p = .009, η^2 = .094); also regarding epistemic quality, there was a significant main effect of VERB (F (1,70) = 9.437, p = .003, η^2 = .119). These results indicate that the VERB prompts improved levels of transactivity as well as epistemic quality, see figures 3 and 4.

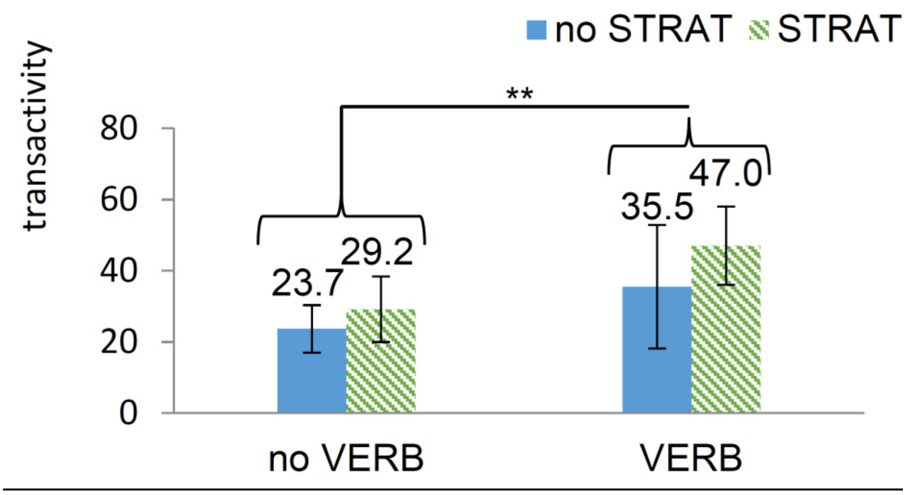


Figure 3. Scores in transactivity; main effect of VERB on transactivity. Error bars represent 95% CIs. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$.

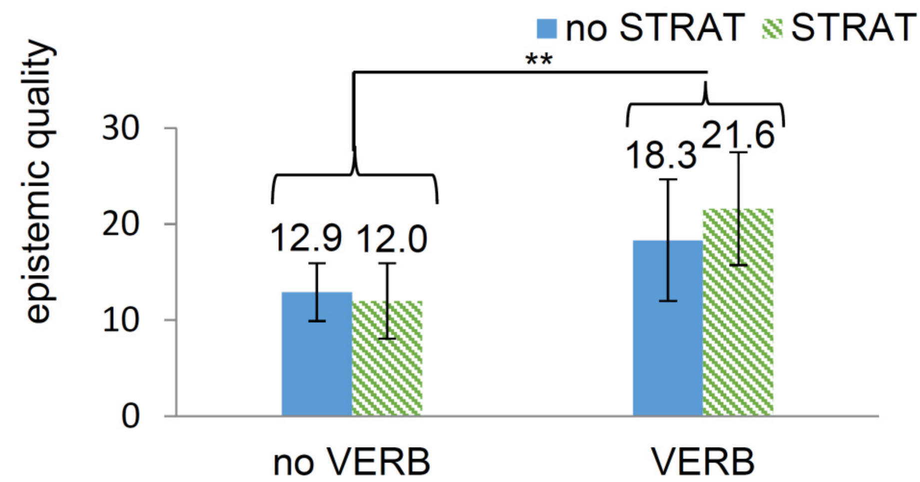


Figure 4. Scores in epistemic quality; main effect of VERB on epistemic quality. Error bars represent 95% CIs. * $p \leq .05$; ** $p \leq .01$; *** $p \leq .001$.

Qualitative analysis of dialogic interactions

To illustrate these findings, we present a qualitative analysis of two exemplary dyads' dialogue. We contrast one dyad from the VERB condition, as a representative for the positive impact of the VERB prompts, with one dyad from the control condition, as a baseline for unprompted "natural" interactions. The dyads were chosen based on their values in three variables: prior knowledge (total points in the pre-test), transactivity score, and epistemic quality score. To ensure comparability, we chose dyads which we found to be "standard" representatives of their condition (values within the first standard deviation from the mean) regarding the aforementioned three variables (see table 7). Both dyads are females only. For this analysis, we name the VERB condition participants “Vanessa” and “Victoria” and the control condition participants “Carolyn” and “Clara”. We performed the analysis on the first analyzed sample (reaction to problem 2:1 and problem solving process of problem 3:1).

Table 7
Selected dyads for the qualitative analysis in comparison to their condition

	Prior knowledge	Transactivity	Epistemic quality
Control: all	10.2 (3.4)	23.7 (14.6)	12.9 (6.6)
Control: Carolyn and Clara	12.5	26.5	12.5
VERB: all	9.6 (3.5)	35.5 (32.6)	18.3 (11.9)
VERB: Vanessa and Victoria	10.0	31.0	20.5

Tables 8 and 9 display excerpts of the selected dyads’ talk during their learning experience with Proportion. Common in both excerpts is a phase of celebrating the success of having solved the previous problem (rows 1 and 2) and a phase of concrete problem solving dialogue which starts in row 9 (VERB) and row 3 (control) respectively. An important distinction between both dyads is the additional “Reflecting” phase (rows 5 to 8) Vanessa and Victoria are engaging in before moving on to solving the next task. Prompted by the pedagogical agent (“Explain to your learning partner…”), they have a phase of explaining each other what was actually going on, taking their time to reflect on *how* they have just solved the task at hand. This additional reflection phase did not take place in the case of Carolyn and Clara who directly moved on with the next problem. These excerpts display how the VERB prompt can facilitate the abstraction from the immediate experience and the verbalization of knowledge.

Table 8
Excerpt from Victoria and Vanessa (VERB)

	Participant	Utterance	Transactivity	Epistemic quality	Phase
1	Victoria	(laughs)	Ext	RegIA	Celebrating
2	Vanessa	Right. That was good.	Ext	RegIA	
3	Victoria	Explain to your learning partner: What …	Acc	RegIA	Prompt
4	Vanessa	… Your learning partner: What did one need to do in order to solve the task?	Acc	RegIA	
5	Victoria	Well, one had to… You had to… First, I went up, then you had to go down.	Ext	TaskIA	Reflecting
6	Vanessa	One had to, wait, one had to find the double.	COC	ContIA	
7	Victoria	Yes	Acc	RegIA	
8	Vanessa	That's double. OK.	Acc	ContIA	
9	Victoria	OK. Not yet.	Acc	RegIA	Problem solving
10	Vanessa	Oh, again now. Hey, stop. You have 3, you need to … still a bit… stop. Not yet. Stop.	Eli	TaskIA	
(…)					

Table 9
Excerpt from Carolyn and Clara (control)

	Participant	Utterance	Transactivity	Epistemic quality	Phase
1	Carolyn	Yes!	Ext	RegIA	Celebrating
2	Clara	We won.	Ext	RegIA	
3	Carolyn	Now, this needs to…	Ext	RegIA	Problem solving
4	Clara	Yes	Acc	RegIA	
5	Carolyn	No, this needs to (mumble) … Wah what are you doing? This is a 3, this needs to be higher than this (mumble)	Eli	TaskIA	
6	Clara	Ah, then we need to lower this one	COC	TaskIA	
(…)					

Hypothesis 4: Effects on learning gains

Hypothesis 4 stated that STRAT and VERB and their combination will increase learning gains, regarding both near and far transfer task types. Two-factorial repeated measures ANOVAs were conducted. Regarding near transfer tasks, the results indicate a general improvement independent from conditions (main effect of point in time): F (1,77) = 11.179, p = .001, η^2 = .13. There was no statistically significant interaction of STRAT × point in time (F (1,77) =.585, p= .447), VERB × point in time (F (1,77) = .061, p = .805) or STRAT × VERB × point in time (F (1,77) = 1.502, p = .224). Regarding far transfer tasks, there was no general improvement from pre- to post-test (no main effect of point in time): F (1,77) = .174, p = .677. There was no statistically significant interaction of STRAT × point in time (F (1,77) = 3.231, p = .076), VERB × point in time (F (1,77) = .112, p = .739) or STRAT × VERB × point in time (F (1,77) = .003, p = .953).

Discussion

Collaborative learning with a multi-touch environment can enable young learners to bodily experience mathematical properties. The learning experiences may be enhanced by different ways of scaffolding. In this study, we enriched the learning environment Proportion with *strategy* prompts and / or *verbalization* prompts to mitigate procedural and conceptual challenges of CSCL (Furberg, 2016). The following paragraphs will first present the outcome of the hypothesis testing, and second synthesize and explain our findings.

Hypothesis 1 claimed a main effect on task focus for the strategy prompts, the verbalization prompts as well as their combination. While the strategy prompts did not have a statistically significant effect on task focus, the verbalization prompts' influence was strong but reverse to our hypothesis: The verbalization prompts actually increased off-task behavior. An interaction effect of the prompts on task focus was not found. Hypothesis 1 needs to be rejected.

Hypothesis 2 predicted more positive emotions caused by the strategy prompts and more negative emotions caused by the verbalization prompts. As hypothesized, the verbalization prompts increased the occurrence of negative emotions with a medium-sized effect. As positive emotions were not significantly affected by strategy prompts, hypothesis 2 is being rejected regarding positive emotions.

Hypothesis 3 predicted higher levels of transactivity and higher epistemic quality caused by the verbalization prompts. The results show that verbalization prompts had positive medium-sized effects on students' dialogues on both, transactivity as well as epistemic quality, supporting hypothesis 3.

Hypothesis 4 claimed that strategy prompts, verbalization prompts, and their combination will increase learning gains, regarding both near and far transfer task types. The results revealed a medium-to-large statistically significant improvement from pre- to post- math test over all conditions for the near transfer tasks. There was no significant improvement regarding the far transfer tasks; also, there was no interaction with any of the prompts. Hypothesis 4 is being rejected.

Overall, the strategy prompts did not affect the measured variables, neither positively nor negatively. Researchers have argued that supportive scaffolding fosters students' motivation by increasing their perceived level of competence (Belland et al., 2013). Our study showed that this might not be universally true and may well be dependent on the type of learning environment. In highly engaging learning environments like ours, learners may not regard strategy prompting as an additional motivator and rather continue solving the tasks. Alike, Jackson et al. (1998) found that students either turned off supportive prompts after a while, or "learned to ignore" (p. 191) them. If learners do not engage with the strategy prompts, they also miss the opportunity to benefit from the provided strategies, which explains their statistically non-significant impact. It may be of interest to focus on developing strategy prompts that are closely tied to the specific misconceptions in a learning domain, e.g. tackling adding and counting strategies in proportional reasoning. These domain-specific prompts may increase learners' motivation to engage with them, because their benefits for solving the problem are more obvious. In the present study, we chose another approach: Similar to Schukajlow et al. (2015), we developed a prompting approach of structuring and breaking down a complex task by inducing an overarching "metacognitive planning strategy" (Schukajlow et al., 2015, p. 1246), rather than aiming at the domain-specific problems. While this showed to be effective in the study by Schukajlow et al. (2015), the effect did not materialize in our study. This may be explained by the fact that our prompts were comparatively subtle, to not interfere too much with the ongoing embodied learning processes, while Schukajlow et al. (2015) provided very detailed and all-encompassing strategy prompting. It is a matter of further empirical research to determine if more intensive prompting has a positive effect by potentially impacting learners' behavior more pronouncedly, or if it has a negative effect by over-scripting learners (Dillenbourg, 2002). Next to motivational factors, learners may also have difficulties in actually understanding and applying the suggested strategies, because of limited cognitive capacities in processing them; productive handling of prompts is a function of time and repeated exposure (Schoenfeld, 1987). Furthermore, one prior study investigated the effects of additional structuring in a computer-*mediated* learning scenario (Baker & Lund, 1997). While they found positive effects of additional structuring on task-focused interactions, the strategy prompts in the present study did not significantly affect task focus or epistemic quality. This brings up the question if computer-mediated communication is more receptive to strategy prompts than face-to-face communication. Finally, Kupers, van Dijk, and van Geert (2017) point out that "scaffolding cannot be planned in advance but instead develops in the lesson itself" (p. 135). Thus, dynamically adapting prompts may be advantageous.

The verbalization prompts had differentiated effects on the learning processes. They enhanced quality of talk, did not affect learning gain, and increased off-task behavior and negative emotions. Our findings of improved quality of talk through prompting for verbalization are in line with findings with adult learners, e.g. showing positive effects of scaffolding argumentation (Weinberger, Stegmann, & Fischer, 2010). Our study shows that verbalization prompts can raise discussion quality also for young learners (see also, Gelmini-Hornsby et al., 2011; Gijlers, Weinberger, van Dijk, Bollen, & van Joolingen, 2013); a very encouraging finding as active knowledge building and sharing is a key for successful CSCL (Roschelle & Teasley, 1995). The exemplary qualitative analysis illustrates this mechanism by showcasing how a dyad, prompted to verbalize, took its time to collaboratively reflect on the learning material, compared to a control condition dyad which did not. It is surprising that this positive effect of verbalization prompts on the quality of talk did not increase learning gains beyond the overall effect. One possible explanation is that verbalization prompts not only raised quality of talk, but also off-task behavior. Sustained task focus is a prerequisite and predictor for learning (Baker & Lund, 1997; Cohen & Lotan, 1995). Consequently, diminished task focus implies less opportunity to continuously engage with the learning environment and, thus, might negatively compensate positive effects of higher-quality verbalizations on learning gains. An alternative explanation is that verbalizations can also corrupt performance, especially in non-verbal tasks and in tasks with hard-to-verbalize cognitive representations (verbal overshadowing effect; Schooler, 2002). Furthermore, social preference might play a role: van Dijk, Gijlers, & Weinberger (2014) found that social aspects are crucial in young learners' collaborative learning, i.e. students who got along well with each other produced better results. In our study, we randomized the dyad formation, so maybe that is why participants did not always collaborate effectively, though there is no indication that this was skewed in favor or disfavor of any one experimental group. Finally, students might not have been able to transfer the verbal abstraction during their collaborative interactions to solving the pen and paper math test problems individually. The effect of verbalization prompts on emotions and task focus are in line with Chen and Law (2016), who found additional question prompts do decrease seventh graders' motivation in a tablet-based learning environment. We reason that the verbalization prompts were in conflict with the learners' subjective needs; they nearly doubled the occurrence of negative emotions and off-task behavior. It seems plausible that these effects are closely interconnected, e.g., increased negative emotions relating to reduced willingness to continuously engage with the task. Learners often do not have a good understanding of what makes their learning effective and so, difficult additional tasks, even if they are designed to help learners, may not be well accepted by the learners, as they disrupt the ongoing immersive activities (Baker & Lund, 1997).

In general, prompts have proven to be useful in supporting learning, and in the present study, we could show how prompting verbalizations can improve the quality of young learners' talk in a multi-touch environment. However, the present study also demonstrates that the type of prompt matters in collaborative learning scenarios, because prompts can also be non-effective or even have adverse effects. There is a thin line between enriching a game-like learning environment in a way that facilitates learning and disengaging students (Deater-Deckard et al., 2014). Further specification of the processes that learners should engage in may allow for more targeted and adaptive scaffolding. Future research may synthesize empirical findings on the effects of different prompting approaches and relate them to an ontology of learning activities. For instance, we may identify further sub-steps of successful task completion and associated learning, and design specific verbalization prompts to support these sub-steps and their associated domain-specific misconceptions.

In conclusion, this study contributes to research on scaffolding collaborative learning by showing how two types of prompts had differentiated effects in a collaborative learning environment with multi-touch devices featuring embodied learning experiences. Beyond overall learning gains, strategy prompting did not influence learning processes or outcomes. Prompting for verbalization increased off-task behavior and negative emotions, but, importantly, supported students dialogue to be more transactive and of higher epistemic quality.

Limitations and future work

This concluding section addresses limitations of the presented study. First, it remains an open question to what extent the conclusions drawn are applicable to students from other countries and cultures. In our study, we included participants from school districts with varying social backgrounds to ascertain representativeness of our sample beyond showcase schools. Evidence is accumulating on the general effects of scaffolding across domains and age groups, but further studies need to replicate and build on our findings to further pin down specific contexts and conditions of scaffolding in multi-touch environments. Second, blinding the video coding was not possible. While the study participants were not aware that there are different experimental conditions or in which condition they were, the video coders unavoidably noticed the experimental condition of the dyad to be coded, because the prompt was part of the coded sample. To achieve objectivity and reliability of results and avoid systematic bias in the data, we made sure that, on one hand, the experimental conditions were equally distributed onto the coders, and on the other hand, there was enough training with a high inter-rater reliability as result. Coder training was framed as so that both, positive and negative outcomes could be expected from any of the prompts. Third, reducing data to selected samples inevitably causes information loss. The presented study's aim was to statistically determine differences in learning processes and outcomes depending on an experimental treatment. To achieve this, a large data set is needed and data reduction methods required. To make sure that the chosen samples are representing the whole data set well, we sampled the participants' learning processes throughout the whole intervention time, thus focusing on behavior during typical problem solving rather than on special cases of initial coordination or final conclusions. Fourth, the aggregation of the individual categories of transactivity and epistemic quality is to some extent a subjective decision, which was taken to allow for quantification and statistical comparison of groups. The aggregation builds on the ordinal scale as it was proposed by Berkowitz and Gibbs (1987). The ordinal scale of categories allows to rank the qualities of the categories. Assigning weights to the subcategories of the scale is necessary to represent the suggested rank order. We apply additional qualitative analyses to validate and illustrate the process analysis.

Future research may focus on how to highlight the value of verbalization prompts to the learners, for example, teachers could practice productive handling of prompts together with the learners before the collaborative phase in a learning environment starts. Furthermore, the long-term effects of scaffolding for verbalization are of interest: Does the quality of talk remain high once the scaffolds have been faded out? In the future, it will be insightful to design and investigate adaptive and personalized prompts, so that quantity, timing and content of the prompts is closely tied to learners' needs. To that end, automatized real-time multi-modal analyses of the learning processes is needed. There is already promising work about automatic facial expression recognition software (Harley, Bouchet, Hussain, Azevedo, & Calvo, 2015) and further following these developments will reveal more insights into collaborative learning and how to support it. Another open challenge is how to orchestrate embodied learning experiences with phases of reflection and abstraction. A possible direction is to disentangle the learning experiences into individual homogeneous phases instead of integrating them. This approach brings up further open issues: First, identifying effective sequencing of the experiences. For example, a recent study found initial hands-on experiencing followed by more traditional exercises to be superior to sequencing these learning opportunities the other way around (Ottmar & Landy, 2017). Second, referencing the different learning experiences, i.e. supporting learners in grasping the relations between hands-on and reflective learning experiences. Taken together, there is a clear need to further research the mechanisms of effective prompting, including the question of how learners appropriate the respective scaffolds (Tchounikine, 2016).

Acknowledgements

The authors would like to thank Jochen Rick for designing and programming the Proportion app and the various student assistants supporting the research project.

Compliance with Ethical Standards

The authors declare that they have no conflict of interest. Informed consent was obtained from all individual participants included in the study. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

Abrahamson, D. (2017). Embodiment and mathematics learning. In K. Peppler (Ed.), *The SAGE encyclopedia of out-of-school learning* (pp. 247-252). Thousand Oaks, CA: SAGE.

Alvarez, C., Brown, C., & Nussbaum, M. (2011). Comparative study of netbooks and tablet PCs for fostering face-to-face collaborative learning. *Computers in Human Behavior*, *27*(2), 834–844.

Baker, M., & Lund, K. (1997). Promoting reflective interactions in a CSCL environment. *Journal of Computer Assisted Learning*, *13*, 175–193.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645.

Belland, B. R., Kim, C., & Hannafin, M. J. (2013). A Framework for Designing Scaffolds That Improve Motivation and Cognition. *Educational Psychologist*, *48*(4), 243–270.

Berkowitz, M. W., & Gibbs, J. C. (1983). Measuring the developmental features of moral discussion. *Merrill-Palmer Quarterly*, *29*(4), 399–410.

Borge, M., & White, B. (2016). Toward the Development of Socio-Metacognitive Expertise: An Approach to Developing Collaborative Competence. *Cognition and Instruction, 34*(4), 323–360.

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: where young children go wrong. *Developmental Psychology*, *44*(5), 1478–1490.

Caballero, D., van Riesen, S. A. N., Álvarez, S., Nussbaum, M., De Jong, T., & Alario-Hoyos, C. (2014). The effects of whole-class interactive instruction with Single Display Groupware for Triangles. *Computers and Education, 70*, 203–211.

Chen, C. H., & Law, V. (2016). Scaffolding individual and collaborative game-based learning in learning performance and intrinsic motivation. *Computers in Human Behavior*, *55*, 1201–1212.

Chi, M. T. H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471–533.

Cohen, E. G., & Lotan, R. A. (1995). Producing equal-status interaction in the heterogeneous classroom. *American Educational Research Journal, 32*(1), 99–120.

Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2008). Gesturing makes learning last. *Cognition*, *106*, 1047–1058.

Danish, J. A., Enyedy, N., Saleh, A., Lee, C., & Andrade, A. (2015). Science through technology enhanced play: Designing to support reflection through play and embodiment. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015: Vol. 1* (pp. 332–339). Gothenburg, Sweden: The International Society of the Learning Sciences.

Davidsen, J., & Ryberg, T. (2017). “This is the size of one meter”: Children’s bodily-material collaboration. *International Journal of Computer-Supported Collaborative Learning, 12*(1), 65–90.

Deater-Deckard, K., El Mallah, S., Chang, M., Evans, M. A., & Norton, A. (2014). Student behavioral engagement during mathematics educational video game instruction with 11-14 year olds. *International Journal of Child-Computer Interaction*, *2*(3), 101–108.

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL. Can we support CSCL?* (pp. 61–91). Heerlen, Netherlands: Open Universiteit Nederland.

Ellis, S., Klahr, D., & Siegler, R. S. (1993). *Effects of feedback and collaboration on changes in children’s use of mathematical rules.* Paper presented at the Meetings of the Society for Research in Child Development, New Orleans.

Falloon, G., & Khoo, E. (2014). Exploring young students' talk in iPad-supported collaborative learning environments. *Computers & Education*, *77*, 13–28.

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*(5), 349–358.

Furberg, A. (2016). Teacher support in computer-supported lab work: bridging the gap between lab experiments and students' conceptual understanding. *International Journal of Computer-Supported Collaborative Learning, 11*(1), 89–113.

Gelman, R., Cohen, M., & Hartnett, P. (1989). To know mathematics is to go beyond thinking that "Fractions aren't numbers." In C. A. Maher, G. A. Goldin, & R. B. Davis (Eds.), *Proceedings of the 11th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education: Vol. 2* (pp. 29–67). New Brunswick, NJ.

Gelmini-Hornsby, G., Ainsworth, S., & O'Malley, C. (2011). Guided reciprocal questioning to support children's collaborative storytelling. *International Journal of Computer-Supported Collaborative Learning*, *6*(4), 577–600.

Gijlers, H., Weinberger, A., van Dijk, A. M., Bollen, L., & van Joolingen, W. (2013). Collaborative drawing on a shared digital canvas in elementary science education: The effects of script and task awareness support. *International Journal of Computer-Supported Collaborative Learning*, *8*(4), 427–453.

Harley, J. M., Bouchet, F., Hussain, M. S., Azevedo, R., & Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior*, *48*, 615–625.

Jackson, S. L., Krajcik, J., & Soloway, E. (1998). The design of guided learner-adaptable scaffolding in interactive learning environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI* (pp. 187–194). Los Angeles, CA: ACM Press.

Jitendra, A. K., Star, J. R., Rodriguez, M., Lindell, M., & Someki, F. (2011). Improving students' proportional thinking using schema-based instruction. *Learning and Instruction*, *21*(6), 731–745.

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology, 55*, 623–655.

King, A. (1990). Enhancing peer interaction and learning in the classroom through reciprocal questioning. *American Educational Research Journal*, *27*(4), 664–687.

Krippendorff, K. (2012). *Content analysis: An introduction to its methodology* (3rd ed.). Thousand Oaks, CA: Sage.

Kupers, E., van Dijk, M., & van Geert, P. (2017). Changing Patterns of Scaffolding and Autonomy During Individual Music Lessons: A Mixed Methods Approach. *Journal of the Learning Sciences, 26*(1), 131–166.

Lehman, B., Matthews, M., D'Mello, S., & Person, N. (2008). What are you feeling? Investigating student affective states during expert human tutoring sessions. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Proceedings of the 9th International Conference on Intelligent Tutoring Systems* (pp. 50-59). Berlin, Heidelberg: Springer.

Linnenbrink, Elizabeth, A., & Pintrich, P. R. (2002). Achievement Goal Theory and Affect: An Asymmetrical Bidirectional Model. *Educational Psychologist, 37*(2), 69–78.

Martin, T., Smith, C. P., Forsgren, N., Aghababyan, A., Janisiewicz, P., & Baker, S. (2015). Learning Fractions by Splitting: Using Learning Analytics to Illuminate the Development of Mathematical Understanding. *The Journal of the Learning Sciences, 24*(4), 593–637.

Mercier, E. M., Higgins, S. E., & da Costa, L. (2014). Different leaders: Emergent organizational and intellectual leadership in children's collaborative learning groups. *International Journal of Computer-Supported Collaborative Learning*, *9*(4), 397–432.

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1999). Early fraction calculation ability. *Developmental Psychology*, *35*(5), 164–174.

Ottmar, E., & Landy, D. (2017). Concreteness Fading of Algebraic Instruction: Effects on Learning. *Journal of the Learning Sciences, 26*(1), 51–78.

Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, *37*(2), 91–106.

Reinholz, D. L., Trninic, D., Howison, M., & Abrahamson, D. (2010). It's not easy being green: Embodied artifacts and the guided emergence of mathematical meaning. In P. Brosnan, D. Erchick, & L. Flevares (Eds.), *Proceedings of the 32nd annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education: Vol. 6* (pp. 1488–1496). Columbus, OH: PME-NA.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *Journal of the Learning Sciences*, *13*(3), 273–304.

Regionalverband Saarbrücken. (2012). *1. Bildungsbericht für den Regionalverband Saarbrücken*. Saarbrücken, Germany: Regionalverband Saarbrücken.

Rick, J. (2012). Proportion: A tablet app for collaborative learning. In *Proceedings of the 11th Annual Interaction Design and Children Conference, IDC* (pp. 316–319). New York, NY: ACM Press.

Rick, J., Kopp, D., Schmitt, L., & Weinberger, A. (2015). Tarzan and Jane Share an iPad. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 1* (Vol. 1, pp. 356–363). Gothenburg, Sweden: The International Society of the Learning Sciences.

Rick, J., Rogers, Y., Haig, C., & Yuill, N. (2009). Learning by doing with shareable interfaces. *Children, Youth and Environments*, *19*(1), 320–341.

Roschelle, J., Rafanan, K., Bhanot, R., Estrella, G., Penuel, B., Nussbaum, M., & Claro, S. (2010). Scaffolding group explanation and feedback with handheld technology: Impact on students' mathematics learning. *Educational Technology Research and Development*, *58*(4), 399–419.

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer Supported Collaborative Learning* (pp. 69–97). Berlin, Germany: Springer.

Sakr, M., Jewitt, C., & Price, S. (2014). The semiotic work of the hands in scientific enquiry. *Classroom Discourse, 5*(1), 51–70.

Sakr, M., Jewitt, C., & Price, S. (2016). Mobile Experiences of Historical Place: A Multimodal Analysis of Emotional Engagement. *Journal of the Learning Sciences, 25*(1), 51–92.

Schneps, M. H., Ruel, J., Sonnert, G., Dussault, M., Griffin, M., & Sadler, P. M. (2014). Conceptualizing astronomical scale: Virtual simulations on handheld tablet computers reverse misconceptions. *Computers & Education*, *70*, 269–280.

Schoenfeld, A. H. (1987). What's all the fuss about metacognition? In A. H. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 189–215). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for Research on Mathematics Teaching and Learning* (pp. 334–370). New York, NY: MacMillan.

Schooler, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Applied Cognitive Psychology*, *16*(8), 989–997.

Schukajlow, S., Kolter, J., & Blum, W. (2015). Scaffolding mathematical modelling with a solution plan. *ZDM Mathematics Education*, *47*(7), 1241–1254.

Sharples, M., Scanlon, E., Ainsworth, S., Anastopoulou, S., Collins, T., Crook, C., … O'Malley, C. (2015). Personal Inquiry: Orchestrating Science Investigations Within and Beyond the Classroom. *The Journal of the Learning Sciences, 24*(2), 308–341.

Sim, G., Cassidy, B., & Read, J. C. (2013). Understanding the fidelity effect when evaluating games with children. In *IDC 2013* (pp. 193–200).

Sim, G., MacFarlane, S., & Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. *Computers & Education*, *46*(3), 235–248. http://doi.org/10.1016/j.compedu.2005.11.021

Tchounikine, P. (2016). Contribution to a theory of CSCL scripts: taking into account the appropriation of scripts by learners. *International Journal of Computer-Supported Collaborative Learning*, *11*(3), 349-369.

Teasley, S. (1997). Talking about reasoning: How important is the peer in peer collaboration? In L. B. Resnick, R. Säljö, C. Pontecorvo, & B. Burge (Eds.), *Discourse, tools and reasoning: Essays on situated cognition* (pp. 361–384). Berlin, Germany: Springer.

van Dijk, A. M., Gijlers, H., & Weinberger, A. (2014). Scripted collaborative drawing in elementary science education. *Instructional Science*, *42*(3), 353–372.

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers & Education*, *46*(1), 71–95.

Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning and Instruction*, *17*(4), 416–426.

Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior*, *26*(4), 506–515.

Wouters, P., van Oostendorp, H., ter Vrugte, J., Vandercruysse, S., de Jong, T., & Elen, J. (2015). The role of curiosity-triggering events in game-based learning for mathematics. In J. Torbeyns, E. Lehtinen, & J. Elen (Eds.), *Describing and studying domain-specific serious games. Advances in game-based learning* (pp. 191–207). New York, NY: Springer.

Study IV:

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets

Lara Johanna Schmitt, Dimitra Tsovaltzi, Armin Weinberger

Abstract

Considerable emphasis has been put on embodied cognition and learning, but this has not yet been applied to questions of collaborative group formation, which typically concentrated on learners' prior characteristics. Little is yet known about the role of heterogeneity in learning processes. Collaborative embodied learning provides the opportunity to study within-group heterogeneity in learning processes, thus, draws on interesting synergies between both research areas. This article explores how differences in bodily and cognitive learning processes within collaborative groups shape their interactions and affect their performance, and how proportional reasoning skills may be developed in a collaborative learning scenario drawing on bodily activities. We report on a study (n = 80) looking at how learners' contributions to solving the task within a dyad diverge and how they are emotionally engaged to different degrees, as indicated by bodily expressions of emotions. We find that within-dyad heterogeneity of cognitive and emotional processes affects their collaborative learning: For example, depending on the heterogeneity of emotional expressivity, learners focus either on high quality talk, or on high bodily performance, but not both. The insights gained in this article inform the development of new ways to structure learners' interactions taking into account their ongoing learning processes.

Keywords: embodied cognition, touch interfaces, heterogeneity, learning processes

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

The theoretical perspective of embodied cognition emphasizes the role of the body and bodily activities for human processing and learning (e.g., Barsalou, 1999; Glenberg, 2010; Wilson, 2002). This perspective may help in supporting young learners' mathematics learning with learning environments that allow for bodily involvement and hands-on engagement with mathematical materials and tasks to foster their understanding (Ottmar & Landy, 2017; Reinholz, Trninic, Howison, & Abrahamson, 2010). In addition, collaborative learning, i.e. multiple learners working on a shared task with the opportunity to exchange ideas and refine each other's understanding (e.g., Dillenbourg, Järvelä, & Fischer, 2009), has shown to enhance seventh graders' performance in mathematics (Webb, 1993). However, the success of collaboration for learning achievement strongly depends on the quality of the interactive processes (Webb, 1991). Learning in collaborative embodied learning environments involves learners on bodily and cognitive modes of interaction. When working face-to-face, learners can engage in dialogue, exchanging their ideas and arguments, they use gestures for illustrating their understanding and interacting with the learning environment, and they express their emotions through their bodies, as well as perceive the bodily activities and expressions of their learning partner (e.g., Davidsen & Ryberg, 2017; Isohätälä, Näykki, Järvelä, & Baker, 2018). In consequence, this richness of co-present embodied learning scenarios implies that to attain an understanding of the learning processes requires a multimodal analysis of not only cognitive aspects of verbal data, but also bodily expression of emotions (see also, Isohätälä et al., 2018).

Prior research identified that learners' heterogeneity within a collaborative group is a crucial factor influencing learning processes and outcomes (e.g., Manske, Hecking, Chounta, Werneburg, & Hoppe, 2015; Webb, 1982). Although learning processes play a major role in collaborative and embodied settings, the majority of research on learner heterogeneity focuses on prior characteristics such as prior knowledge, or differences in demographical background such as gender. Hence, a largely unexplored factor is the role of heterogeneous learning processes. Are there differential effects when learners within a group engage in rather homogenous or heterogeneous processes during their learning activities? For example, with respect to the degree of verbal elaboration of concepts, or the extent to which they express their emotions? Cognitive and bodily aspects of a learning process can become very relevant for small-group learning activities which are based on bodily activities. To advance our understanding of how learners acquire knowledge and gain a better mathematical understanding, research needs to address not only individual learning, but also the role of learning processes in context (i.e., in a collaborative embodied learning setting). Particularly in collaborative learning, where multiple learners come together with homogeneous or heterogeneous learning processes, questions of learner heterogeneity become relevant. Collaborative embodied learning with touch screen tablets may be a productive scenario for young learners' acquisition of proportional reasoning skills (Schmitt & Weinberger, 2019), a typically challenging domain in mathematics (Mix, Levine, & Huttenlocher, 1999), and provide a rich context for researching the role of bodily and cognitive process heterogeneity in collaborative groups.

In this article, we present a study to shed more light on the conditions of children's collaborative embodied learning. Specifically, we look at the effects of heterogeneous collaborative processes in an embodied learning environment. To this end, this article will introduce relevant related work on collaborative learning with embodied learning technologies, embodied cognition and bodily expression of emotions, and the role of group heterogeneity. We aim at enhancing our understanding on how learners' multifaceted learning processes unfold and interact while they collaboratively construct meaning with the help of multi-touch technology.

Learning in collaborative embodied learning environments

In collaborative learning settings, participants are invited to discuss and refine concepts interactively with the goal of co-constructing knowledge and of achieving shared understanding, i.e. greater knowledge convergence, as a result (Weinberger, Stegmann, & Fischer, 2007; Roschelle, 1992). Collaborative learning and CSCL (computer supported collaborative learning) have been widely investigated, and evidence is accumulating regarding the learning processes involved, the instructional design that supports effective learning processes, and the analytic and methodological frameworks for their analysis, which carry theoretical assumptions about collaborative learning (e.g., Dillenbourg et al., 2009; Weinberger & Fischer, 2006). Although a lot of prior research on collaborative learning focuses on adults co-constructing knowledge, e.g. through online argumentation (Tsovaltzi, Judele, Puhl, & Weinberger, 2015; Kollar, Fischer, & Slotta, 2007), also children and teenagers as technology users and collaborative learners have received considerable attention in research (e.g., Roschelle & Teasley, 1995; Stahl, 2006). Collaborative learning scenarios for primary and secondary education span computer-mediated as well as co-present scenarios with vast opportunities for different technological tools to support their collaborative learning, like tabletops, tablets, desktop PCs, social networks, etc. (Schmitt & Weinberger, 2018). In mathematics education, instruction traditionally was geared towards symbol manipulation, disregarding learners' informal and intuitive knowledge and learning processes (Resnick, 1989). In general, while collaborative learning has often focused on processes of joint reasoning, recent research also includes more and more aspects of bodily and emotional involvement in learning, and their combination. For example, the international CSCL conference in 2019 had a focus on '4E learning' – learning as a process which is embodied, enactive, extended, and embedded (Lund et al., 2019).

In the present study, we investigate children's collaborative learning with a touch interface affording for bodily activities. Therefore, the following paragraphs will introduce the concept of 'embodied cognition', and how the design of learning environments may acknowledge the role of the body for learning. Moreover, we will discuss how embodied learning experiences may need to be combined with opportunities for abstraction and reflection, and how cognitive and bodily / emotional processes in such an embodied environment develop.

Embodied cognition for collaborative learning

The embodied cognition perspective

The perspective of embodied cognition points at the role of bodily experiences as non-reducible resources for how we perceive, process and make use of conceptual knowledge, emotions and social information (Niedenthal, Barsalou, Winkielman, Krauth-Gruber, & Ric, 2005). Embodied cognition approaches assume that rather than having an isolated and private mind that processes information like a computer, humans are involved with their full body in understanding and interacting with the environment (Spackman & Yanchar, 2014). Hence, cognition necessarily builds on sensorimotor processing, also in seemingly purely cognitive activities that are not directly related to any immediate interaction with the physical surroundings (Wilson, 2002).

In embodied cognition, a distinction is made between online and offline embodiment. Online embodiment refers to the real-time embodied processing of information while a stimulus (for example, an emotional stimulus) is present, and offline embodiment refers to mental simulations of past experiences in the absence of this stimulus, and without actually carrying out the motor response (Niedenthal et al., 2005; Wilson, 2002). Analogously, Black, Segal, Vitale, and Fadjo (2012) distinguish physical from imagined embodiment in their Instructional Embodiment Framework, i.e. the actual bodily vs. purely mental simulation. Embodied cognition has been influential on the design of learning environments, in the sense that active bodily experiences, e.g. touch, object manipulation, and bodily movements are included (e.g., Black et al., 2012; Reinholz et al., 2010). We use the term 'embodied learning environment' to refer to learning environments that explicitly integrate online embodiment, i.e. bodily activity. For example, touch interfaces involve bodily activity by directly manipulating objects (Black et al., 2012; Ottmar & Landy, 2017).

Tablets for embodied learning

Touch interfaces, such as tablets, can support learning. In the domain of mathematics learning, 6-7 year old children demonstrated higher efficiency, and improved use of strategies when learning with a tablet compared to learning with a traditional interface (a laptop with a mouse) (Segal, 2011). In collaborative learning, touch interfaces are typically fun to use (Clayphan, Collins, Kay, Slawitschka, & Horder, 2018). Some findings indicate a decrease in negatively charged perceptions of dominance by group members and an increase of bodily awareness of other group members in such an embodied learning environment, as well as beneficial effects on physical participation, negotiation of physical interactions, and collaboration (Jakobsen & Hornbæk, 2016; Marshall, Hornecker, Morris, Dalton, & Rogers, 2008). Touch interfaces also offer potential of embodied learning for cognitive processes and outcomes. For example, prior research on children's collaborative learning showed that using a shared tabletop fostered 10-year-old children's mathematics learning compared to a similar paper-based activity; specifically, the tabletop condition fostered flexibility in using the mathematical concept (Mercier & Higgins, 2013). Children aged 10-12 learned successfully about photosynthesis by collaboratively drawing this scientific phenomenon on a shared tablet (van Dijk, Gijlers, & Weinberger, 2014). Even dyads of 5-year-old children used tablets successfully while engaging in on-task activities almost all of the time (Falloon & Khoo, 2014).

According to the Instructional Embodiment Framework (Black et al., 2012), embodied learning experiences should be actively included into formal teaching settings, while the instruction can focus on facilitating either physical or imagined embodiment. Black et al. (2012) hypothesize that a good instructional method would be to start with physical embodiment to fully experience the concept, then move on with imagined embodiment, before engaging with more abstract transfer tasks. However, rather than choosing between 'bodily' and 'cognitive' learning experiences, it may make sense to enrich bodily involvement with cognitive reflection in discourse as negotiating meaning is a developing process. Advanced approaches of educational design include cycles of practical hands-on learning activities with cycles of reflective activities (e.g., Ottmar & Landy, 2017). Bodily and verbal actions (and their interaction) are thereby central in learners' joint construction of understanding (Gomoll, Hmelo-Silver, Tolar, Šabanovic, & Francisco, 2017). Bodily activities can accompany the verbal discussion (Gomoll et al., 2017) and bodily expressivity may help learners negotiate and grasp important concepts together as they become aware of each other's bodily / emotional states, represent each other's sensorimotor activities and interact productively in collaborative learning environments (Rick, Kopp, Schmitt, & Weinberger, 2015; Davidsen & Ryberg, 2017; Rick, 2012).

Supporting embodied learning

Interacting in embodied learning environments requires some level of guidance and opportunities for reflection (e.g., Schneps et al., 2014). This might be even more the case in collaborative settings. Support for collaborative argumentative learning through collaboration scripts has often revealed positive results (e.g., Vogel et al., 2013). Little is known, however, about the effects of different forms of support for embodied learning environments, e.g. for strategically maneuvering oneself in embodied interactions, and to what extent and how such a support may impact bodily and cognitive learning processes and learning outcomes.

In a previous study, we designed verbalization and strategy prompts to support learners' joint reasoning in the embodied learning environment 'Proportion' (Schmitt & Weinberger, 2019): Verbalization prompts elicited explanations to foster abstraction of embodied experiences, and strategy prompts structured learners' task progress to make them carefully plan their collaborative embodied activities. Independent from the experimental conditions, there were significant learning gains from pre- to post-test. The verbalization prompts improved the quality of dialogue regarding epistemic quality and transactivity, i.e. how learners use and connect relevant concepts, and the extent to which they refer to their partner's contributions (see Weinberger & Fischer, 2006). They also strikingly increased negative emotions and off-task behavior, which seem to have prohibited further learning gains (Schmitt & Weinberger, 2019).

More work needs to be done in understanding the ongoing learning processes in an embodied learning environment, and how interactions on the bodily dimension may relate to interactions on the cognitive dimension, and how cognitive and bodily interactions relate to outcomes. For example, analyzing how the experience and the expression of emotions relate to learners' epistemic quality in their verbalizations may generate valuable insights for understanding collaborative embodied learning. Emotional experiences and their bodily expression are crucial factors for collaborative activities. Therefore, the scope of the present article is to identify and untangle cognitive and bodily conditions of successful collaborative learning and zoom in on the effects of cognitive and bodily learning processes as they develop after receiving verbalization prompts.

Collaborative learning and emotions

Collaborative learners may be strongly involved emotionally regarding oneself, the task, their performance, the context, and interactions with peers (Järvenoja & Järvelä, 2005). Emotional responses can be negatively experienced, like anxiety, or positive experienced, like satisfaction (Järvenoja & Järvelä, 2005), and shape groups' interactions and learning by guiding learners' attention and affecting co-regulation (Polo, Lund, Plantin, & Niccolai, 2016). The intensity of emotions that learners experience during their collaboration can shape the nature of their discussion, i.e. if learners engage in cumulative, disputational, or exploratory talk (see Mercer, 1996): For example, intense emotions are associated with disputational talk, where learners rather argue and try to get their point across instead of coming to a consensus (Polo et al., 2016). In contrast, low levels of emotions are associated with cumulative talk, where learners reach a consensus but with limited critical discussion (Polo et al., 2016). In order to engage in cognitive processes of argumentation, learners need to take care of the group emotional processes. The challenge is to create a good atmosphere without overly prioritizing the good climate to the extent that a critical discussion would not be possible anymore and the learners simply agree on a quick but superficial consensus (Isohätälä et al., 2018).

Processing of emotions is embodied (Niedenthal et al., 2005), i.e. affective and bodily states are closely related and mutually influence each other: Emotions are expressed through the body, and bodily states influence perceived emotions (Barsalou, 2008; Ekman, 1992; Strack, Martin, & Stepper, 1988). In a collaborative learning situation, emotional processes are represented not only through learners' verbal reasoning and argumentation, but also become evident through their bodily behaviors, e.g. through gestures (Isohätälä et al., 2018). These non-verbal features of the interactions may help understanding learners' emotions and their relation to learning processes (Isohätälä et al., 2018). It can also help uncovering the manifold of interactions which take place between learning partners and between learners and technology (Davidsen & Ryberg, 2017). For example, learners do not only talk, but make use of their fingers and arms to demonstrate understanding, store information, regulate access to a device or information, or express disagreement (Davidsen & Ryberg, 2017). Especially in an embodied learning environment, not only the interaction (e.g., touch) with technology itself, but also learners' bodily activities "in the zone in-between" become relevant (Davidsen & Ryberg, 2017, p. 67). This 'zone' refers to the open space between the group members, or between the members and the screen, and allows for movements, gestures, or touches which are crucial for collaboration, coordination and learning, for example simulating a touch interaction in midair to explain something to the learning partner (Davidsen & Ryberg, 2017).

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

There are inter-individual differences regarding the extent to which learners express their emotions in observable behavior (gestures), or oppress/conceal them (Järvenoja & Järvelä, 2005; see also Ekman, 1992). As our study by design involves bodily movements and engages learners, we assume that different levels of bodily expression of emotions will be observable.

To understand the multi-faceted interactions unfolding in a collaborative and embodied learning environment better, in this study, we measure and analyze learning processes on multiple dimensions. The dimensions include the bodily expression of emotions, cognitive processes operationalized as task focused behavior and the quality of dialogue (transactivity and epistemic quality, e.g. Weinberger & Fischer, 2006), and performance, i.e. efficiency within the learning environment, knowledge outcomes and knowledge convergence. The following section will elaborate on how heterogeneity of ongoing processes may be relevant for collaborative learning.

Heterogeneity in collaborative learning

Let us imagine two young learners (around 10 years) sharing a tablet interface to solve mathematical problems by manipulating the touch screen with their fingers and trying out different configurations of the digital objects on the screen. These learners may not be used to regularly using tablet apps for learning purposes, and on top of that, to use one shared device with a classmate. A wide range of emotions and emotional expressions may develop during this activity. The learners may feel happy when they complete a level, feel discouraged when a task seems unsolvable, or approach their learning partner with shared pride and express that by clapping into their hands, or become angry when the partner behaves in a certain way. Characteristics of the individual learners may fundamentally change the effects of any learning intervention or specific designs of learning environment. For example, in individual learning settings, participants' cognitive resources, emotional states, and personality traits may affect multi-media learning (Knörzer, Brünken, & Park, 2016). In collaborative learning settings, these prior learner characteristics can be distributed homogeneously or heterogeneously within a group.

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Heterogeneity in prior characteristics

Group member heterogeneity can occur on many levels; the research focus often was on differences in prior knowledge, demographical background, or attitudes (e.g., Harrison, Price, Gavin, & Florey, 2002; ter Vrugte et al., 2015). For example, it was found that heterogeneous groups (based on domain knowledge and motivation) showed better performance, higher learning gain, and smoother interaction than homogeneous groups (Manske et al., 2015). In contrast, homogeneous low-ability students tend to produce a lot of off-task behavior, as they may become frustrated by the perceived difficulty of the task (Webb, 1982). Regarding cognitive variables such as the kind of prior knowledge, there seems to be a consensus in research that, overall, heterogeneous group formation is advantageous for collaborative learning (ter Vrugte et al., 2015), because it can foster productive cognitive conflict (Jorczak, 2011). The collaborative information processing model (CIP) "predicts that heterogeneous groups will learn better than homogeneous groups" and "that students must be sufficiently heterogeneous in their knowledge to increase the probability that expressed information will diverge and conflict" (Jorczak, 2011, p. 216-217). Cognitive heterogeneity may stimulate active construction of a shared understanding (Gomoll et al., 2017), and promote supportive behavior in mixed ability groups (Webb, 1982). However, there are also findings that prior knowledge heterogeneity within a group may result in dominance with the more able student dominating the interactions, particularly when the learning group is in competition with other groups (ter Vrugte et al., 2015). This dominant behavior manifests in high participation by the more able student and in low participation by the less able student, and ultimately impedes collaborative processes (ter Vrugte et al., 2015).

Group member heterogeneity may also shape emotional variables. For example, university students with more or less favorable goal structures (regarding performance, well-being and learning in collaborative learning) influenced each other's emotional appraisal regarding the collaborative learning activity, depending on the level of group heterogeneity (Wosnitza & Volet, 2012). Even though no differences in the appraisal were found between groups prior to the activity, groups with a homogeneous goal profile reinforced each other, i.e. in groups with a homogeneously high goal profile appraisal increased, whereas in groups with a homogeneously low goal profile, appraisal decreased as a result of the collaborative learning activity (Wosnitza & Volet, 2012). Heterogeneous goal profile groups developed rather positive appraisals, which means that students with initially less favorable collaborative goal structures can profit from collaborating with students that have a more favorable goal structure (Wosnitza & Volet, 2012). Other examples of relevant dimensions for learner heterogeneity are social preference, learning style, or productivity. Grouping young students (10-12 years) according to their social preference increased learning gains and the quality of the joint product (van Dijk et al., 2014). Heterogeneity of learning styles within a group did not influence achievement, but had minimal effects on satisfaction favoring heterogeneous groups (Lehman, 2007). In the context of collaborative work, there is evidence that forming groups of heterogeneous workers (in terms of individual productivity rates) positively affects average productivity of the group compared homogeneous groups (Hamilton, Nickerson, & Owan, 2003).

Potential benefits of cognitive heterogeneity interact with other learner characteristics and the quality of the group processes. For example, low achievers may profit to a different extent than high achievers (Cheng, Lam, & Chan, 2008), and the characteristics of other learners shape perceptions of one’s own characteristics (frame-of-reference hypothesis, respectively big fish in a small pond effect, Marsh & Parker, 1984). Moreover, high-quality group processes can balance out effects of group formation (Cheng et al., 2008). Last, group formation aspects also interact with the learning environment design: For example, heterogeneity of competence levels can be advantageous, but only when the environment allows for hypothesis testing (Asterhan, Schwarz, & Cohen-Eliyahu, 2013). Also the effects of different experimental conditions (e.g., touch vs. mouse input) become less pronounced when the group is homogeneous with respect to talk levels (Marshall et al., 2008).

In conclusion, levels of group homogeneity and heterogeneity are of great relevance for successful collaboration, for example, group formation may shape the sort of emotional processes learners develop (Polo et al., 2016). However, the mechanisms of group heterogeneity influencing emotional processes during collaborative learning are still not well understood. Despite many positive findings in favor of heterogeneous groups, the overall pattern is not consistent (Cheng et al., 2008). Schwarz and Linchevski (2007, p. 512) state that “while initial cognitions are relevant to further peer collaboration, the number of possibilities is immense and it is impossible in general to predict whether confronting cognitions will lead to argumentation and learning" – a problem exacerbated by the fact that most studies on group heterogeneity consider only prior learner characteristics.

It is worthwhile not to depend on prior measures of user traits, but to consider the heterogeneity of processes that evolve in the time frame of the collaboration. Next to questions of orchestrating interactions, the topics of emotional awareness and group formation have been identified as the main challenges in CSCL research, which require thorough investigations to inform the development of effective learning environments (Reis et al., 2018; Tsovaltzi et al., 2019). Further, there is a need to better understand the potentially positive effects of heterogeneity in learning groups and under which conditions they unfold (Wosnitza & Volet, 2012). Cheng et al. (2018) stated that productive group processes (operationalized as positive interdependence, individual accountability, equal participation and social skills) were more relevant than questions of heterogeneity in preconditions regarding the reported collective efficacy of the group. However, prior research on heterogeneity in (bodily and cognitive) processes is sparse. The following paragraphs will synthesize the few available findings on process heterogeneity.

Heterogeneity in cognitive and bodily processes

Regarding heterogeneous *bodily* processes, prior research (Rick et al., 2015) highlighted – in a case study – how large heterogeneity in emotional valences and task focus was turned into a productive co-regulatory process of 'taming' and 'stimulating', resulting in remarkable learning gains. Thus, beyond verbal processes, bodily strategies are crucial in socially coordinating collaborative learning, e.g. by helping to manage the group's focus (Jamil et al., 2017), and bodily states have an impact on interactions and mutual understanding between humans (Glenberg, 2010). There are differences in physical strategies cross-culturally (Jamil et al., 2017) and interpersonally (Järvenoja & Järvelä, 2005), e.g. in the extent learners express their emotions physically (gestures). Due to differing physical experiences that individual humans have, they develop individually different representations of objects or concepts, however they are also able to share these with another through (verbal) interaction ('shared embodiment', Barsalou, 1999). If differences in bodily processes become salient, they may result in higher awareness of the learning partner and deliberate attempts to explain the individual embodied representations to each other to reach a common proposed bodily action. As bodily expressions and gesturing are crucial in describing spatial elements of a situation accurately (Rauscher, Krauss, & Chen, 1996), explaining one's representation of an embodied task may foster learning. However, heterogeneous bodily processes within a collaborative group may also require additional effort to regulate each other, which can distract learners from the main learning task.

In the present study, we will test the hypotheses that the degree of heterogeneity in bodily processes within a group will affect the group's cognitive processes (hypothesis 1), and the group's performance (hypothesis 2). The effects may be positive or negative: On one hand, bodily heterogeneity may raise awareness of these bodily divergences and possibly trigger cognitive activities to resolve them, like explaining diverging bodily representations or strategies to each other and sustaining constant focus towards the task (enhanced cognitive processes). Perceived bodily divergences may also be turned into productive embodied interactions, which help them to make progress within the app and are conducive to learning (enhanced performance). On the other hand, bodily heterogeneity may result in distraction and greater need for co-regulation and managing these bodily divergences, therefore hindering productive learning processes (diminished cognitive processes), and meaningful learning or engagement with the tasks of the app (diminished performance).

Regarding heterogeneity in *cognitive* processes, it has been found that heterogeneous ways of thinking and expressing ideas, or diverging opinions in groups, can productively enrich discussions and foster learning by providing multiple perspectives (Curşeu, Schruijer, & Boroş, 2012; Rosebery, Ogonowski, DiSchino, & Warren, 2010; Tsovaltzi, Weinberger, Scheuer, Dragon, & McLaren, 2013). One prior study showed that inducing socio-cognitive conflicts in groups through awareness tools fostered high quality argumentation and knowledge acquisition (Puhl, Tsovaltzi, & Weinberger, 2015). In contrast, cognitive divergences can also result in socio-emotional process losses in groups, e.g. in increased social conflicts and worsened group climate (Curşeu et al., 2012). Based on these mixed results we conclude that differences in cognitive processes may become salient through learners' explanations and have differentiated effects, positive as well as negative ones.

In the present study, we will test the hypothesis that the degree of heterogeneity in cognitive processes within a group affects the group's performance (hypothesis 3). Prior research indicated that diverging cognitive processes may have positive or negative effects, such as enriching discussions, but also process losses (Curşeu et al., 2012). On one hand, heterogeneous cognitive processes may be advantageous, for example, a learner who is capable of engaging in high-quality verbalizations may profit from explaining their understanding to a less able partner, and a learner who has difficulties in explaining concepts or engaging in verbal abstraction may benefit from a more able peer who can scaffold and model these processes. On the other hand, diverging cognitive processes may reduce common ground and a common 'language' for interacting productively with each other, and therefore hinder learning.

In sum, we expect heterogeneous cognitive and bodily processes alike to impact learning processes and outcomes. We argue that, for a better understanding of collaborative embodied learning, one needs to look beyond heterogeneity in prior learner characteristics, and shift the focus to learners' heterogeneity in ongoing learning processes. Prior research on heterogeneity of learner traits has produced mixed results, and prior research on heterogeneity in processes is still sparse. Thus, a statistical analysis of process heterogeneity will enable us to gain more insights into the effects of process heterogeneity, to what extent process heterogeneity on different dimensions plays a role, and if process heterogeneity is rather productive or detrimental to learning. Because both cognitive and bodily processes and process heterogeneity on these dimensions are relevant in collaborative learning, we analyze how bodily processes relate to cognitive processes, and to what extent heterogeneity of cognitive and bodily processes influences performance. Before presenting the methods and results of this study, the following section will shortly introduce the learning domain (proportional reasoning) and the learning environment (the 'Proportion' app).

Proportional reasoning and learning with 'Proportion'

Proportional reasoning

Proportional reasoning is defined as ''reasoning with ratios, rates, and percentages'' (Jitendra, Star, Rodriguez, Lindell, & Someki, 2011, p. 731), or ''understanding the multiplicative relationships between rational quantities'' (Boyer, Levine, & Huttenlocher, 2008, p. 1478). Proportional reasoning is a foundation for higher-order mathematical reasoning, but often associated with misconceptions and difficulties (Boyer et al., 2008). By requiring to go beyond adding and counting strategies, and to understand multiplicative relationships (Boyer et al., 2008), proportional reasoning involves a fundamentally new way of manipulating numerical quantities for primary school children.

Involving the body and providing hands-on embodied learning experiences may help children to grasp proportional relations (Reinholz et al., 2010). A hands-on learning experience may be realized with a learning intervention with manipulatives, building on children's intuitive knowledge (Fujimura, 2001). Physical manipulation of tangible objects, in this case magnets on a magnetic board, significantly improved fourth graders' performance in a proportional reasoning test (Fujimura, 2001). More recent approaches integrate technologies to help children to embody proportional relations, for example the Mathematical Image Trainer (Reinholz et al., 2010). An advantage of technology-supported embodied learning environments is the technology's ability to provide fast and accurate feedback.

The 'Proportion' learning environment

'Proportion' (Rick, 2012) is a tablet app to foster proportional reasoning of fourth graders aged around 10 years. The basic goal in Proportion is to manipulate two vertical bars that have a number associated each to be in the right relation to each other. An owl acts as a pedagogical agent and provides feedback. Once the correct relation between the bars is set, the text notification "correct" appears on screen, the owl makes a cheerful sound, and the next task (the next numerical relation) appears. The tasks are becoming increasingly difficult as learners make progress, e.g. the easier tasks involve setting the relation between two integers (e.g., 2 : 3), the harder tasks involve setting the relationship between two fractions with the same (2/5 : 4/5) or even different (2/5 : 3/7) denominator. In total, Proportion provides 216 or these tasks organized in 21 levels. See figure 1 for a screenshot of a typical Proportion task (level 1).

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

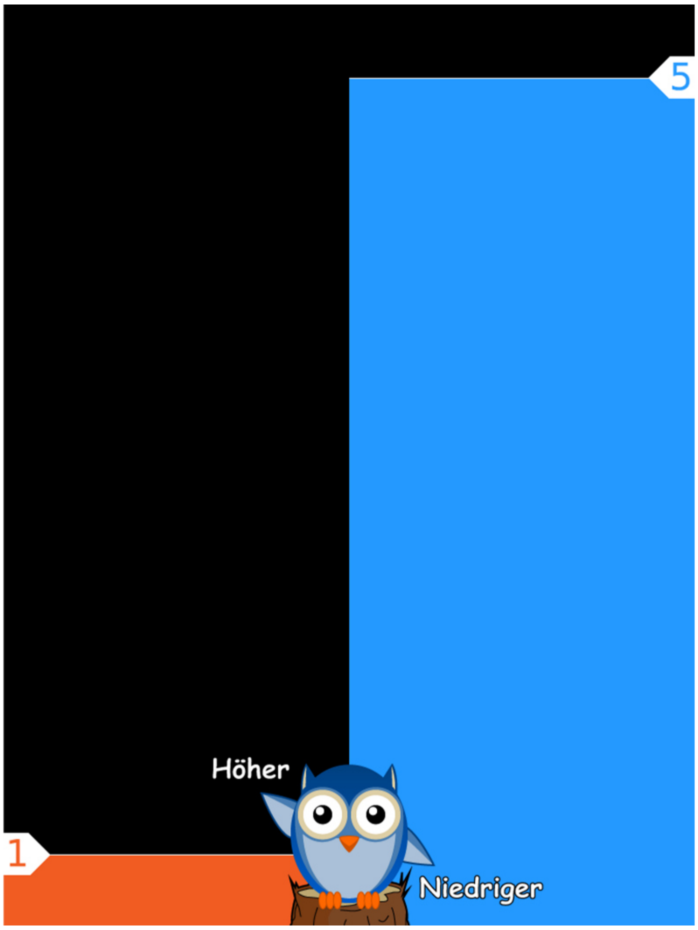


Figure 1. Screenshot of Proportion with the task 1:5. The owl provides feedback by indicating "higher" (höher) and "lower" (niedriger) when the learners do not solve the task after one minute.

The Proportion app is designed for two co-present collaborative learners (Rick, 2012). With the learners sharing and manipulating the same relatively small representation (the tablet), they are sitting right next to each other and can interact directly and closely with each other and the application, supporting awareness of the learning partner's bodily activities (Rick, 2012), speech, and emotional expressions. The touch interface affords active physical manipulations through hand movements for actively embodying proportional relations, targeting *Direct Physical* (Black et al., 2012), respectively *Online* embodiment (Niedenthal et al., 2005).

In conclusion, we assume that collaborative learning with Proportion incorporates two elements of embodied learning. First, the bodily movement required by the app influences the online sensorimotor representation of proportions. Second, the tablet app's affordance for bodily movement may facilitate bodily interactions between the learners allowing for free bodily expression of emotions. For example, a momentum of happiness or frustration is expressed and recognized through accompanying bodily expressions (e.g., clapping into the hands), making the learning experience socially embedded and embodied.

Research question and hypotheses

Based on the theoretical considerations and empirical findings presented, the goal of the present study is to shed more light on the role of bodily and cognitive processes, and their heterogeneity, as possibly determining conditions of co-present collaborative embodied learning. As previous research has revealed partially surprising results, e.g. contra-intuitive effects of verbalization prompts (Schmitt & Weinberger, 2019), a closer look at learners' multi-faceted processes and how they are distributed and unfold within dyads is needed. The hypotheses are tested with the sub-sample that received the verbalization prompt from Schmitt and Weinberger (2019). The present analysis is exploring new aspects of group heterogeneity, specifically heterogeneity in learning *processes*. As the empirical findings on process heterogeneity aspects so far are not allowing clear predictions regarding their effects, the hypotheses which are leading the process analysis are tentative and non-directional, but nevertheless crucial to enable a statistical analysis.

In our study, learners' bodily processes are reflected through their bodily expression of emotions. Beyond bodily processes, we are interested in cognitive processes relevant to productive collaborative learning, namely task focus, i.e. learners' persistence of engaging with the learning environment, and the quality of their dialogue, regarding their ability to build up on each other (transactivity) and their ability to abstract from the concrete task towards the underlying concepts (epistemic quality) (see Weinberger & Fischer, 2006). In addition, we also regard learners' performance, both their bodily efficiency within the app, i.e. how many problems they solve, as well as the outcome in a knowledge test after the learning phase (knowledge outcome) and the degree to which learners within one group converge on their knowledge as a result of working together (knowledge convergence). In the previous chapter on heterogeneity in collaborative learning, we constituted the hypotheses on the effects of process heterogeneity in collaborative embodied learning. In short, we assume the following non-directional hypotheses:

(1) Heterogeneity in bodily processes affects the quality of cognitive processes.
(2) Heterogeneity in bodily processes affects the groups' performance.
(3) Heterogeneity in cognitive processes affects the groups' performance.

Methods

Sample

We analyzed n = 80 participants (mean age: 10 years and 5 months (SD = 6.9 months); 51.3 % male). The participants were fourth graders from local primary schools in Germany. The dyad formation in this study was randomized, to avoid uninformed assumptions about how prior learner characteristics influence processes beyond purely cognitive ones.

Experimental procedure

Experiments took place inside the respective schools and replaced two regular school lessons (90 min). Trained experimenters welcomed the participants and provided an introduction to the upcoming activities. Then, participants individually filled in a first survey and a prior knowledge test (10 minutes). In the following learning phase, participants collaboratively used the Proportion app for 40 min. We video-taped this learning phase (one camera plus microphone per dyad). Last, participants individually filled in a second survey and a post-test (10 minutes). In addition to the baseline Proportion environment, verbalization prompts requested learners to explain, summarize, and generalize (see Schmitt & Weinberger, 2019).

Materials and variables

For capturing multi-modal process variables, we transcribed and video coded specified samples of the video material throughout the entire learning period of 40 minutes. The chosen sample units (every second task from every second level) allow for measuring the variables during typical problems that evoke emotions and observable cognitive processes (solving a problem). Detailed coding schemes for the qualitative analysis of the process variables *negative emotions*, *positive emotions*, *task focus*, *transactivity*, and *epistemic quality* are published in Schmitt and Weinberger (2019); in the present article, we give a summarized overview on how we captured these variables. Inter-rater reliability values for these individual variables were obtained for the entire original sample presented in Schmitt and Weinberger (2019) using Krippendorff's Alpha with $.77 \leq \alpha \leq .92$.

As we were interested in the effects of heterogeneity in learning process variables, we needed to calculate an indicator of heterogeneity within the dyads. To categorize a dyad as homogeneous vs. heterogeneous with respect to a process variable, we utilized the Coefficient of Variation (CV; Weinberger et al., 2007) for this respective variable. The CV is a relative measure of homogeneity within a group; it is determined by first calculating a group's mean and standard deviation for the respective variable, and then dividing the standard deviation by the mean (Weinberger et al., 2007). The resulting CV value is useful for comparing different levels of heterogeneity across groups. Using a median split, we categorized the dyads with a CV within the range of the lower ~50% of the sample as homogeneous groups, and the dyads with a CV within the range of the upper ~50% of the sample as heterogeneous groups. Dependent variables in our analyses where aggregated on dyad level, as we consider the dyad, and not the individual, as the unit of analysis.

For the present analysis, we regard three main aspects of relevant learner variables: Bodily processes, cognitive processes, and performance.

Regarding *bodily processes*, we focus on bodily expression of emotions. The coding scheme for measuring emotions focused on instances where the body was visibly involved in expressing the emotion, i.e. through body movements (Schmitt & Weinberger, 2019; see also, Lehman, Matthews, D'Mello, & Person, 2008; Sim, MacFarlane, & Read, 2006). In the emotions coding scheme presented in Schmitt and Weinberger (2019), positive and negative emotions were regarded as separate variables (e.g., Frijda, 1988; Linnenbrink & Pintrich, 2002), and measured by coding and counting the frequency of bodily expressed positive vs. negative emotions per analyzed problem in the video data. For instance, positive emotions were thumbs-up or clapping into one's hand, instances for negative emotions were threatening the tablet or a dismissive hand gesture (Schmitt & Weinberger, 2019). However, also seemingly 'negative' physical interactions, like physical disagreement, contribute to successful interaction (Fleck et al., 2009). Therefore, for the present analysis, we merged positive and negative emotion indicators into one combined variable comprising the overall amount of bodily expression of emotions. See figure 2 for an example of a dyad with heterogeneous bodily processes.

Figure 2. Two learners with heterogeneous bodily processes. The learner on the left is currently expressing a (positive) emotion after the dyad successfully solved a problem in Proportion, while the learner on the right does not.

Regarding *cognitive processes*, we focus on *off-task behavior*, *transactivity*, and *epistemic quality*. Off-task behavior, respectively task focus, can be operationalized as observable (non-verbal) behavior (Deater-Deckard, El Mallah, Chang, Evans, & Norton, 2014). We measured this variable through coding and counting the instances of off-task behavior per analyzed problem in the video data (Schmitt & Weinberger, 2019). For instance, looking around in the classroom or into the camera was counted as an instance of off-task behavior. Next to off-task behavior, we measured two variables originating from learners' dialogue: *transactivity* and *epistemic quality*. Transactivity indicates the extent to which learners refer to their partner's contributions, and epistemic quality refers to the content of learners' talk, i.e. their ability to adequately incorporate the relevant concepts of the learning domain (Weinberger & Fischer, 2006). In Schmitt and Weinberger (2019), we categorized learners' utterances as different levels of transactivity and epistemic quality with an adapted coding scheme from Weinberger and Fischer (2006). Then, we aggregated and weighted the categories to two global scores, according to their relative quality (Schmitt & Weinberger, 2019).

Regarding *performance*, we focus on *knowledge outcomes*, *knowledge convergence*, and *efficiency*. Knowledge was measured with a math test comprising near and far transfer tasks (Schmitt & Weinberger, 2019). Knowledge outcomes refer to the number of points in the post-test. Knowledge convergence refers to the degree of heterogeneity (CV) in the post-test, i.e. the lower the heterogeneity, the higher the knowledge convergence between learners. Efficiency is operationalized as the total number of solved problems within the Proportion app during the 40 minutes of the learning phase and can be determined through log files.

Overview on analysis

The three hypotheses were tested with MANOVAs and follow-up univariate analyses. For hypothesis 1, three outliers needed to be removed to meet the assumptions of the multivariate analysis. The MANOVAs test the effects of bodily or cognitive group heterogeneity on a set of related dependent variables – either cognitive processes or performance. As will be seen in the results, the MANOVA outcomes of hypotheses 1 and 2 led to a mediation hypothesis to explain these results. The mediation model is tested using the "bootstrap confidence interval for the indirect effect" approach reported in Hayes and Rockwood (2017, p. 44). This approach focuses on the indirect effect of predictor X on outcome Y to determine a possible mediation; if its bootstrapped confidence interval does not include zero (i.e., if it is entirely in either the minus or the plus range), the mediation model is supported (Hayes & Rockwood, 2017). Finally, we present a qualitative case study to illustrate and further explain the quantitative effects.

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Results

First, we were interested in determining general learning gains as measured by pre-post-test changes. The data shows that, overall, there were small but significant improvements in near transfer tasks (t (39) = -2.980, p = .005, d = .257), but not in far transfer tasks (p > .05). Next, we present the results for the three hypotheses of the study.

Hypothesis 1: Heterogeneity in bodily processes affects the quality of cognitive processes

There is an overall large significant effect of heterogeneity of bodily processes on cognitive processes: F (3, 26) = 4.732, p = .009, *Pillai's Trace* = .353, ηp^2 = .353. Specifically, a large negative effect on transactivity is marginally significant: F (1, 28) = 3.796, p = .061, ηp^2 = .119. Groups who are homogenous in terms of bodily processes are more transactive by trend. There are no significant effects of heterogeneity of bodily processes on epistemic quality, or task focus (p > .05). These results provide partial support for hypothesis 1, i.e. heterogeneity in bodily processes affects cognitive processes. Specifically, bodily homogeneity tends to increase the degree of transactivity.

Hypothesis 2: Heterogeneity in bodily processes affects the groups' performance.

There is an overall large significant effect of heterogeneity of bodily processes on performance variables: F (3, 31) = 3.858, p = .019, *Pillai's Trace* = .272, ηp^2 = .272. Specifically, there is a large significant positive effect on efficiency: F (1, 33) = 9.179, p = .005, ηp^2 = .218. Groups that are heterogeneous with regard to bodily processes solve more problems. There is no effect for knowledge outcomes or knowledge convergence (p > .05). In sum, these results provide partial support for hypothesis 2, i.e. heterogeneity in bodily processes affects performance. Specifically, bodily heterogeneity increases the number of problems the participants manage to solve in the learning environment.

Hypothesis 3: Heterogeneity in cognitive processes affects the groups' performance.

There is an overall large significant effect of heterogeneity of epistemic quality on performance variables: F (3, 30) = 3.096, p = .042, *Pillai's Trace* = .236, ηp^2 = .236. In particular, a medium negative effect on knowledge outcomes just misses significance (F (1, 32) = 4.025, p = .053, ηp^2 = .112), and a large negative effect on knowledge convergence is significant with F (1, 32) = 8.120, p = .008, ηp^2 = .202. Groups that are homogeneous with regard to epistemic quality tend to show higher knowledge outcomes and arrive at more similar knowledge scores within groups. However, heterogeneity of epistemic quality does not affect efficiency, and there are no overall significant effects of either heterogeneity of transactivity or heterogeneity of task focus on performance variables ($p > .05$). In sum, these results provide partial support for hypothesis 3, i.e. heterogeneity in cognitive processes affects performance. Specifically, epistemic homogeneity fosters, by trend, learners' knowledge outcomes in the post-test and group knowledge convergence.

Mediation analysis

Based on the results of hypotheses 1 and 2, we pose a mediation hypothesis, namely that the degree of transactivity may be a mediating variable for the effect of heterogeneity of bodily processes on efficiency in the app (see Figure 3). The bootstrapped (5000 samples) unstandardized indirect effect of X (heterogeneity of bodily processes) on Y (efficiency) equals -8.242. The 95% confidence interval ranges from -17.452 (BootLLCI) to -.5160 (BootULCI), therefore, we can conclude that the effect of heterogeneity of bodily processes on efficiency is indeed mediated by the degree of transactivity.

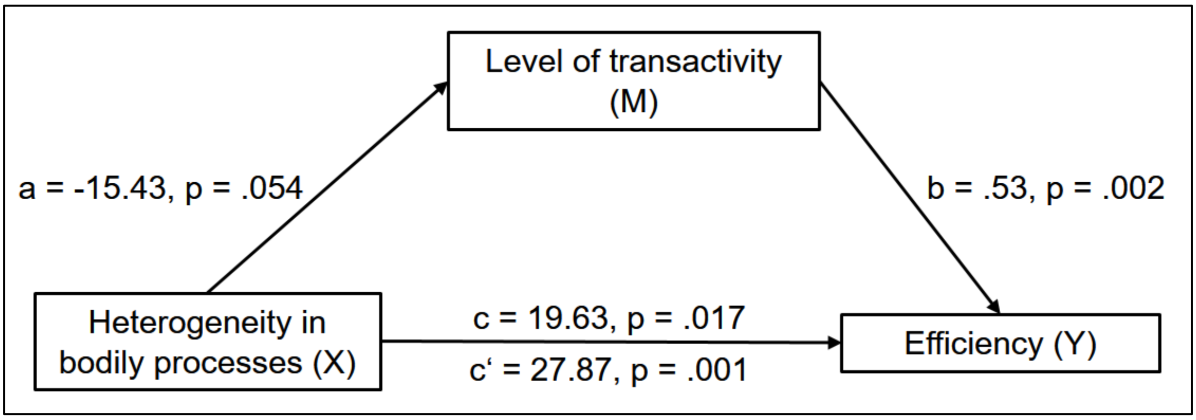


Figure 3. Standardized regression coefficients for the relationship between heterogeneity of bodily processes and efficiency, mediated by level of transactivity (c = total effect, c‘ = direct effect).

Qualitative case analysis

To illustrate the interplay of learners' bodily and cognitive processes based on their group heterogeneity, we present a case study of one dyad. The detailed observation of a selected case is suitable for gaining in-depth insights (Randolph, 2008), and may enhance our understanding of the mechanisms of bodily heterogeneous processes. We explore the multimodal processes of one dyad, which is constituted of two 10 years old learners (one male, one female), who we name Willy and Fiona in this analysis. Based on their bodily expressions of emotions (as recorded by the structured video analysis), they were classified as a bodily heterogeneous dyad: While Willy appears rather emotionally reserved (he showed only one positive emotion during our coded sample, and zero negative ones), Fiona visibly and regularly expresses her emotions through her body (two positive and four negative emotions during our coded sample). As a group, they are very efficient in working with the app and solve 128 problems in 40 minutes, which is more than one standard deviation above the average value of M = 93,75 (SD = 23,78) problems solved. However, they do not manage to engage in very transactive talk: Their combined transactivity score is 23.50, which is rather low compared to the average transactivity score of M = 42.3 (SD = 28.37). We look at their learning processes in the middle of the intervention time (around minute 20). By this time, the learners are familiar with the Proportion tasks and the overall setting and they gained some experience with each other's ways of working, talking, or expressing emotions. Figure

4 displays key screenshots from the excerpt. Table 1 displays the detailed verbal and bodily processes as they unfold.

At minute 19:50, the dyad is confronted with a new task: The target proportion 10:30. At first, both of them play around with the bars a bit, until Willy exclaims “Hey, what are you doing?”. This is the starting point for a more goal-directed behavior: Fiona states “I am looking for the 10”, and Willy starts measuring the bar with his fingers, i.e., he uses his body to estimate the size of a proportion (screenshot 1). Upon finding a solution for himself, he instructs Fiona “To this line here”. Fiona slowly moves the bar towards the position that Willy is pointing at, until he stops her by exclaiming “Stop!”. Fiona releases her hand and throws them in the air dramatically, as in an anticipation of a positive result (screenshot 2). However, the problem is not solved yet. The dyad keeps on refining their solution. Willy verbally instructs Fiona through saying “more down” without touching the tablet himself, and Fiona executes the order and brings the bar more down (screenshot 3). When Fiona goes down too far, he intervenes, both verbally and bodily, with saying “no” in a harsh tone and tries to wipe away her hand. They continue working in a way that Fiona physically engages with the task and moves the bar, while Willy instructs and refines her movements in a mostly calm tone “more up…no… you are too much up… there”. While exclaiming “there”, Willy points his finger to the position that he thinks is correct and leaves the finger there until Fiona moves the bar to this position (screenshot 4); he further instructs here by explaining “exactly on here”. The problem is still not solved and the owl provides the feedback “almost” (almost correct). At this point, Willy moves the tablet closer to himself and says “Let me try it”. A little conflict emerges as both learner try to get physical access to the tablet (screenshot 5). Willy now moves the blue bar up very carefully, while Fiona seems to become desperate and puts her head down on the desk (screenshot 6). Finally, the solution is correct, and Willy seems relieved and says “so”. He moves the tablet back towards his partner and Fiona willingly joins in solving the next problem (screenshot 7).

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Table 1

The unfolding verbal and bodily processes of two bodily heterogeneous learners during the solution of a 'Proportion' task.

Time	Verbal interaction	Bodily interaction	Comment
19:50 – 19:52		Willy: Playing with the orange bar Fiona: Playing with the blue bar	No clear goal or solution strategy recognizable
19:53 – 19:56	Willy: *Hey, what are you doing?* Fiona: *I am looking for the 10.*		Starting point for goal-directed behavior
19:57 – 20:00		Willy: Measuring with his fingers Fiona: Slowly moves the blue bar down	Embodied approach to task solution (measuring, manipulating) [screenshot 1]
20:01	Willy: *To this line here.*	Willy: Points at a position on the tablet that Fiona should move to	Willy guides Fiona verbally and bodily
20:01 – 20:02		Willy: Keeps on pointing with his fingers Fiona: Moves the blue bar to the position that Willy points at	Bodily instruction through pointing gesture
20:03	Willy: *Stop!*		Verbal co-regulation
20:04 – 20:05		Fiona: Touches the owl and dramatically throws hands in the air	Fiona shows bodily anticipation of success [screenshot 2]
20:06	Solution is not yet correct; owl gives feedback: "almost"		
20:06	Willy: *More down.* Fiona: *Oh!*		Willy guides Fiona verbally [screenshot 3]

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Time	Verbal interaction	Bodily interaction	Comment
20:06 – 20:08		Fiona: Slowly moves the blue bar down	Fiona appropriates the task bodily
20:09	Willy: *No!*	Willy: Tries to wipe away Fiona's hand	Willy guides Fiona bodily and verbally
20:10 – 20:12	Willy: *More up…no… you are too much up… there!*	Fiona: Moves the blue bar up	Willy guides Fiona verbally
20:13 – 20:17	Willy: *Exactly on here!*	Willy: Points at a position on the tablet that Fiona should move to Fiona: Fiona: Slowly moves the blue bar down	Willy guides Fiona bodily and verbally and Fiona bodily executes the instructions [screenshot 4]
20:18	Fiona: [Mumbles something]	Fiona: Touches the owl	
20:18	Solution is still not correct; owl gives feedback: "almost"		
20:19 – 20:24	Willy: *Let me try it!*	Willy: Pulls the tablet to himself; physically dominates access to the tablet; carefully moves the blue bar up Fiona: First tries to get access back, then puts the head down on the table in frustration	A little conflict regarding physical access emerges between the learners; Fiona bodily disengages the task and seems to express despair [screenshot 5 and 6]
20:24	The problem is solved		
20:24	Willy: *So!*		
20:25		Willy: Puts the tablet back in the middle Fiona: Physically joins back in	The physical conflict is resolved through physical action (giving back access) [screenshot 7]

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

This short except of the video data illustrates how a bodily heterogeneous group is involved in solving a task. Rather than relying on language and verbal discussions about proportions, this dyad is highly engaged with their whole bodies. For example, they use pointing behaviors to guide the learning partner, they measure the bars with their fingers, or they express their emotions related to their task progress. Conflicts emerge through physical interactions, but are also resolved through physical interactions, e.g. fighting for or giving access to the tablet (screenshots 5, 6, and 7). It becomes evident that Fiona is emotionally more expressive than Willy, as seen by her dramatic hand movement (screenshot 2) or her frustration (screenshot 6), while Willy remains calm and less expressive. While the learners do closely coordinate each other, e.g. Willy often instructs and guides Fiona, the interactions do not require much language, but are based on bodily interactions like measuring and pointing. The analysis of this dyad shows how proportional concepts do not need to be explained through words, but through actions. Concretely, while one could discuss that 10:30 is basically the same as 1:3, Willy rather guides Fiona's bodily movements to a certain position on the tablet. Together, the dyad displays productive bodily processes and strategies and solves a difficult task without many words. As shown by the mediation analysis, this way of bodily solving tasks can be a productive mode of interaction.

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

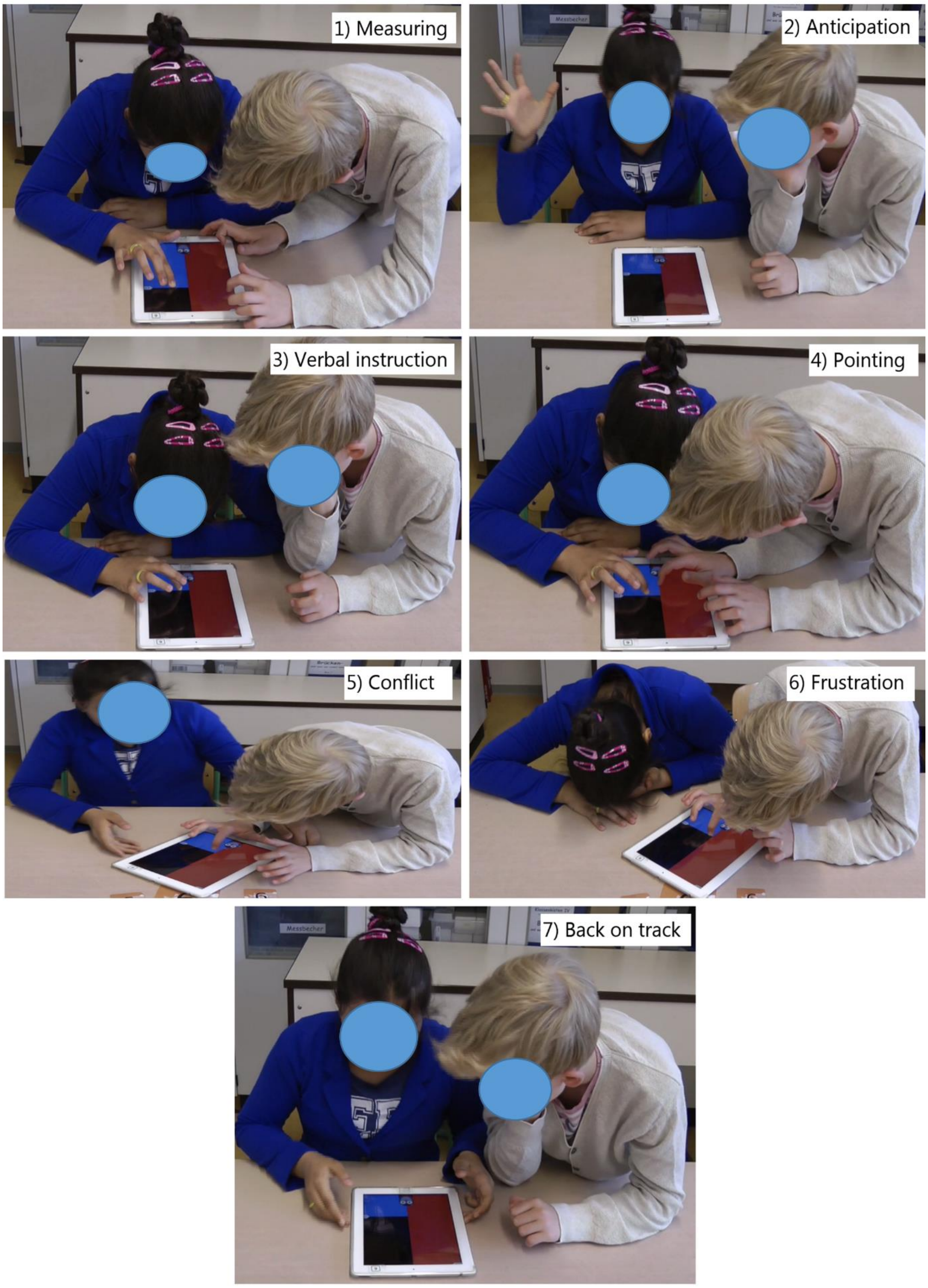


Figure 4. Screenshots of the task progress of two bodily heterogeneous learners during the solution of a 'Proportion' task.

Discussion

In this study, we investigate the effects of bodily and cognitive process heterogeneity in collaborative embodied learning. Summarized, the results provide evidence that both heterogeneous as well as homogeneous processes can be differentially conducive for learning. Homogeneous bodily processes tended to facilitate cognitive processes. Heterogeneous bodily processes enhanced bodily efficiency within the learning environment. Moreover, homogeneous cognitive processes were conducive to knowledge convergence and to knowledge outcomes, by trend. A mediation analysis showed that transactivity mediates the effect of heterogeneity in bodily processes on efficiency. Drawing on a qualitative case analysis, we further illustrate how learners may largely rely on diverse bodily activities – rather than verbal interactions – to solve mathematical problems.

In hypothesis 1, we assumed that heterogeneity in bodily processes would influence learners' cognitive processes. Perceived bodily heterogeneity may produce productive verbal activities to resolve these differences, but may also have negative effects by requiring a higher effort for regulating each other. In our study, we found that homogeneous bodily processes fostered, by trend, the degree of transactivity (a cognitive processes variable). This finding may be explained by assuming that homogeneity in bodily expression decreases time and effort needed to co-regulate diverging emotions and facilitates shared cognitive co-processing between partners. But what happened in bodily heterogeneous groups? Assuming that learners are aware of their bodily divergences, they still may have difficulties in articulating these perceived divergences verbally and resolving them through discourse. Large differences in bodily expressing emotions may not readily result in a group climate which affords elaborate discussions (see Isohätälä et al., 2018). Moreover, the learners in this study were prompted to reflect and abstract towards the underlying concepts of proportional reasoning, in addition to the bodily activities required by the app. As indicated by our analysis, this abstraction through verbalizations was especially prominent for learners with homogeneous bodily processes: Learners who share similar degrees of bodily expressing emotions are more likely, by trend, to react transactively to each other's verbal contributions than learners who have diverging levels of bodily expressivity. For learners with heterogeneous bodily processes, however, prompting for deeper cognitive processes may rather interfere with their bodily activities. This finding highlights that cognitive prompts, e.g. for self-explanations (Chi, Leeuw, Chiu, & Lavancher (1994), can be counterproductive in learning environments where learners take bodily actions in the environment. While it may be worthwhile to develop new scaffolding methods that help bodily heterogeneous learners to engage in productive transactive verbalizations, transactivity in discussion may not be the only way of resolving heterogeneity in bodily processes. Possibly, an emphasis on bodily interactions and bodily task engagement may be a better way for turning bodily process heterogeneity into productive learning processes.

This line of thinking is substantiated by the results regarding hypothesis 2 (heterogeneous bodily processes affect performance). We found that heterogeneous bodily processes increase efficiency, i.e. the amount of solved problems within the embodied learning environment. Potentially, bodily heterogeneity highlights the bodily aspects of the collaborative task solving process, supporting full embodied immersion and construction of deep-level mental representations which are crucial for sustainable learning (Barsalou, 1999). However, we did not find effects of heterogeneous bodily processes on knowledge outcomes and convergence. Bodily learning may not readily transfer to purely cognitive tasks that traditionally populate post-tests. Other structuring forms might be needed to achieve that kind of cognitive abstraction towards formal symbols in case of heterogeneous bodily processes.

Finally, we assumed that heterogeneous cognitive processes affect performance (hypothesis 3). Diverging knowledge and perspectives by theory may foster productive cognitive conflict (e.g., Puhl et al., 2015; Curşeu et al., 2012), and therefore, may be beneficial for learning, but also negatively affect the group climate (Curşeu et al., 2012). In our data, we found that homogeneous cognitive processes fostered knowledge outcomes (by trend) and knowledge convergence. This result may be interpreted by assuming that homogeneous groups talk on the same level and about the same concepts, facilitating making sense of each other's contributions, which would foster understanding of the domain. With homogeneous cognitive processes, it is also plausible that learners converge in their knowledge. Moreover, this aligns with findings, that in epistemically homogeneous groups, the partner's information may be incorporated to a greater extent, as perceived similarity increases people's willingness to accept information from another person (Marks, Copland, Loh, Sunstein, & Sharot, 2018).

Heterogeneous cognitive processes, however, affected performance negatively. Plausibly, one cannot expect high performance from learners who have to cope with and regulate their heterogeneous cognitive processes, and, at the same time, need to bodily interact with a learning environment. Moreover, in groups with heterogeneous cognitive processes, the more able student may have also dominated the partner, hindering productive collaboration (ter Vrugte et al., 2015). Our finding of beneficial homogeneous cognitive processes therefore serves as an extension to prior research on cognitive heterogeneity in groups, and may help to explain previous inconclusive results.

Based on these findings, we assumed that, depending on the groups' constitution with regard to bodily processes, the groups choose different ways of interacting: Homogeneous groups focus on discussing and learning by means of high quality dialogue, and heterogeneous groups focus on interacting intensely with the app and learning by doing. This assumption was supported by the mediation analysis, which indicated that the degree of transactivity mediates the relationship between bodily heterogeneity and effectivity: the higher the bodily heterogeneity, the lower the transactivity, and the higher the efficiency. Concretely, if learners in a group have homogeneous bodily processes, they are more verbally transactive, and therefore less bodily efficient, i.e. solve less problems in the learning environment. In contrast, if learners in a group have heterogeneous bodily processes, they are less verbally transactive, and therefore more bodily efficient, i.e. solve more problems in the app. These findings point to future research directions for embodied learning: Salient diverging bodily processes may take away the focus from the verbalizations, decreasing transactivity, but enhance bodily activities with the app, increasing efficiency. These differing approaches may be characterized as "talkers" vs. "makers". The case analysis of Willy and Fiona further illustrates the way of interacting in the case of a bodily heterogeneous dyad: Rather than discussing the Proportion tasks verbally, they strongly engage the task with their bodies, and utilize their fingers to point at something, guide their learning partner, or regulating access. This focus on bodily behaviour during task solution may be the result of the perceived bodily divergence within the group, as substantiated by the mediation analysis. However, both approaches, jointly reflecting on underlying principles (the "talkers") as well as focusing on hands-on training with the app (the "makers"), appear to be conducive to learning about proportions to the same degree, as we did not find an effect of bodily process heterogeneity on knowledge outcomes.

Conclusion

Technology enables learners from diverging backgrounds, with diverging opinions, prerequisites, behaviours, and abilities to learn together. Further looking into effects of different levels of heterogeneities may help us making collaborative learning more enjoyable and effective by grouping learners in ways that support them in understanding each other and acquire new knowledge. Approaching mathematical principles with touch interfaces that incorporate active bodily manipulations rather than abstract symbol manipulation shows to be a promising way for improving children's proportional reasoning. Overall, the learners significantly improved from pre-test to post-test in near transfer tasks. Embodied learning is a social issue because the body enables the communication to the outside world and, thence, to learning partners with the opportunity to share individual embodied representations (see Barsalou, 1999). This is then central for collaborative learning which is based on interactions with the others to co-construct knowledge. In this article, we analyzed process heterogeneity on multiple dimensions, bodily and cognitive, and showed that learner interactions are influenced both by cognitive and by bodily processes, and that these processes also interact.

Contrary to some of the evidence that cognitive heterogeneity in collaborative groups fosters learning, mostly tested with one-dimensional learner characteristics, we found evidence for both homogeneous as well as heterogeneous processes to be advantageous depending on which dimension of embodied learning was considered. Heterogeneous bodily processes furthered bodily performance, and homogeneous bodily processes, by trend, increased transactivity. Homogeneous cognitive processes fostered, by trend, knowledge outcomes and knowledge convergence. These results emphasize the assumptions that we need to look beyond heterogeneity prior to learning, and consider also process heterogeneity during collaborative interactions. Moreover, as we found differentiated results depending on the process dimension of interest (cognitive or bodily), it is crucial to untangle different process dimensions during collaborative learning.

This study may inform the design of touch environments for learning. First, evidence is accumulating that it is crucial to consider the interplay between bodily and cognitive activities. Incorporating bodily activities into learning environments may be necessary to foster individual understanding, and to take up an embodied perspective when analyzing learning environments, i.e. to regard learners' bodily expressions, including emotional expressions, and interactions with a device and with the learning partner, during the learning process. Furthermore, group formation aspects can change learners' interactions, for example, in this present study, heterogeneous groups with respect to bodily processes were more productive within the learning environment. Moreover, learners may benefit from prompts that foster their ability to abstract from the concrete problems encountered in the learning environment to the underlying concepts (Schmitt & Weinberger, 2019), but may potentially also benefit from prompts that highlight process divergences of learners and provides recommendations how to utilize them productively. In the present study, we found that homogeneity regarding epistemic quality within a group is crucial for learning success (hypothesis 3). Therefore, learners may also benefit from prompts that suggest to them to listen carefully to their partner and to adopt a common language, in order to promote homogeneously high epistemic talk within groups and therefore foster learning outcomes. The results in general seem to speak for the need for congruent supports in case of heterogeneity at any level: Cognitive heterogeneity may require cognitive support, bodily heterogeneity may require support for the bodily expressivity within a group.

Limitations and future research

We focused on a deep process analysis to explore the relationships between different learning processes in an embodied learning environment. As the present article addresses new aspects of group heterogeneity, focusing on learners' cognitive and bodily processes, there was only very limited prior research available. Therefore, our hypotheses were non-directional, and the results supported them only partially. While we provide new insights, these insights are explorative in nature and need further experimental testing in future research. For example, future studies may experimentally vary the group formation to arrive at more pronounced heterogeneity effects. Borderline non-significant results may be a reason of small statistical power. The statistical power was diminished, because we analyzed on dyad level and further divided the sample in heterogeneous vs. homogeneous groups via a median split, reducing the sample size subjected to the statistical analysis. The rather high effect sizes, however, indicate substantial differences and effects, which may become clearly significant with a greater sample size. These limitations may be addressed with future studies that will explicitly form groups based on differing levels of, for example, bodily expression of emotions, and work with larger sample sizes. Future research may also give more insights regarding the generalizability of the results, i.e. to what extent the present results translate to other samples and environments, or how they develop depending on instructional approaches like prompting.

Our empirical findings on process heterogeneity may be utilized for informing innovative ways of group formation practices. Though it may be challenging to spontaneously form groups based on processes that may not be known in advance and cannot be measured as easily and efficiently as prior knowledge, current research is working on technological solutions for assessing learner characteristics and processes. One promising development into that direction is work on group formation based on automatic text mining approaches (Erkens, Bodemer, & Hoppe, 2016). We are confident that the analyses provided in this article can inform and inspire future research to look deeper into the effects of heterogeneous learning processes within groups of learners. Further developing theories of learning is a prerequisite and important step to advance new opportunities of analyzing learners' interactions, e.g. measuring transactivity with machine learning algorithms (Gweon, Jain, Mcdonough, Raj, & Rosé, 2013).

Acknowledgements: The authors would like to thank Jochen Rick for designing and programming the Proportion app and the various student assistants supporting the research project.

References

Asterhan, C. S. C., Schwarz, B. B., & Cohen-Eliyahu, N. (2013). Conceptual change in proportional reasoning: Effects of collaboration, own / partner reasoning level and hypothesis testing. In *Proceedings of the 35th Annual Meeting of the Cognitive Science Society* (pp. 1780–1785).

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577–660. doi:10.1017/S0140525X99002149

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645. doi:10.1146/annurev.psych.59.103006.093639

Black, J. B., Segal, A., Vitale, J., & Fadjo, C. (2012). Embodied cognition and learning environment design. In D. Jonassen & S. Land (Eds.), *Theoretical foundations of learning environments* (2nd ed., pp. 198–223). New York: Routledge. doi:10.4324/9780203813799

Boyer, T. W., Levine, S. C., & Huttenlocher, J. (2008). Development of proportional reasoning: Where young children go wrong. *Developmental Psychology, 44*(5), 1478–1490. doi:10.1037/a0013110

Cheng, R. W.-Y., Lam, S.-F., & Chan, J. C.-Y. (2008). When high achievers and low achievers work in the same group: The roles of group heterogeneity and processes in project-based learning. *British Journal of Educational Psychology, 78*(2), 205–221. doi:10.1348/000709907X218160

Chi, M. T. H., De Leeuw, N., Chiu, M.-H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*(3), 439–477. doi:10.1207/s15516709cog1803_3

Clayphan, A., Collins, A., Kay, J., Slawitschka, N., & Horder, J. (2018). Comparing a single-touch whiteboard and a multi-touch tabletop for collaboration in school museum visits. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1), 6:1-6:23. doi:10.1145/3191738

Curşeu, P. L., Schruijer, S. G. L., & Boroş, S. (2012). Socially rejected while cognitively successful? The impact of minority dissent on groups' cognitive complexity. British *Journal of Social Psychology, 51*(4), 570–582. doi:10.1111/j.2044-8309.2011.02023.x

Davidsen, J., & Ryberg, T. (2017). "This is the size of one meter": Children's bodily-material collaboration. *International Journal of Computer-Supported Collaborative Learning, 12*(1), 65–90. doi:10.1007/s11412-017-9248-8

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Deater-Deckard, K., El Mallah, S., Chang, M., Evans, M. A., & Norton, A. (2014). Student behavioral engagement during mathematics educational video game instruction with 11-14 year olds. *International Journal of Child-Computer Interaction, 2*(3), 101–108. doi:10.1016/j.ijcci.2014.08.001

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning: from design to orchestration. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-enhanced learning: Principles and products* (pp. 3–19). Dordrecht, Netherlands: Springer Netherlands. doi:10.1007/978-1-4020-9827-7_1

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion, 6*(3-4), 169–200. doi:10.1080/02699939208411068

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

Falloon, G., & Khoo, E. (2014). Exploring young students' talk in iPad-supported collaborative learning environments. *Computers & Education, 77*, 13–28. doi:10.1016/j.compedu.2014.04.008

Fujimura, N. (2001). Facilitating children's proportional reasoning: A model of reasoning processes and effects of intervention on strategy change. *Journal of Educational Psychology, 93*(3), 589–603. doi:10.1037/0022-0663.93.3.589

Fleck, R., Rogers, Y., Yuill, N., Marshall, P., Carr, A., Rick, J., & Bonnett, V. (2009). Actions speak loudly with words: Unpacking collaboration around the table. *ITS '09: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, Banff, Canada*, 189–196. New York, NY: ACM. doi:10.1145/1731903.1731939

Frijda, N. H. (1988). The laws of emotion. *American Psychologist, 43*(5), 349–358. doi:10.5465/AMR.2007.25275690

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(4), 586–596. doi:10.1002/ejsp.670

Gomoll, A. S., Hmelo-Silver, C. E., Tolar, E., Šabanovic, S., & Francisco, M. (2017). Moving apart and coming together: Discourse, engagement, and deep learning. *Journal of Educational Technology and Society, 20*(4), 219–232.

Gweon, G., Jain, M., McDonough, J., Raj, B., & Rosé, C. P. (2013). Measuring prevalence of other-oriented transactive contributions using an automated measure of speech style accommodation. *Computer-Supported Collaborative Learning, 8*(2), 245–265. doi:10.1007/s11412-013-9172-5

Hamilton, B. H., Nickerson, J. A., & Owan, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy, 111*(3), 465–497. doi:10.1086/374182

Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface- and deep-level diversity on group functioning. *Academy of Management Journal, 45*(5), 1029–1045. doi:10.5465/3069328

Hayes, A. F., & Rockwood, N. J. (2017). Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour Research and Therapy, 98*, 39–57. doi:10.1016/j.brat.2016.11.001

Isohätälä, J., Näykki, P., Järvelä, S., & Baker, M. J. (2018). Striking a balance: Socio-emotional processes during argumentation in collaborative learning interaction. *Learning, Culture and Social Interaction, 16*, 1–19. doi:10.1016/j.lcsi.2017.09.003

Jakobsen, M. R., & Hornbæk, K. (2016). Negotiating for Space?: Collaborative Work Using a Wall Display with Mouse and Touch Input. *CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose*, CA, 2050–2061. New York, NY: ACM. doi:10.1145/2858036.2858158

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Jamil, I., Montero, C. S., Perry, M., O'Hara, K., Karnik, A., Pihlainen, K., … Subramanian, S. (2017). Collaborating around digital tabletops: Children's physical strategies from India, the UK and Finland. *ACM Transactions on Computer-Human Interaction, 24*(3), 23:1-23:30. doi:10.1145/3058551

Jitendra, A. K., Star, J. R., Rodriguez, M., Lindell, M., & Someki, F. (2011). Improving students' proportional thinking using schema-based instruction. *Learning and Instruction, 21*(6), 731–745. doi:10.1016/j.learninstruc.2011.04.002

Jorczak, R. L. (2011). An information processing perspective on divergence and convergence in collaborative learning. *International Journal of Computer-Supported Collaborative Learning, 6*(2), 207–221. doi:10.1007/s11412-010-9104-6

Järvenoja, H., & Järvelä, S. (2005). How students describe the sources of their emotional and motivational experiences during the learning process: A qualitative approach. *Learning and Instruction, 15*(5), 465–480. doi:10.1016/j.learninstruc.2005.07.012

Knörzer, L., Brünken, R., & Park, B. (2016). Emotions and multimedia learning: the moderating role of learner characteristics. *Journal of Computer Assisted Learning, 32*(6), 618–631. doi:10.1111/jcal.12158

Kollar, I., Fischer, F., & Slotta, J. D. (2007). Internal and external scripts in computer-supported collaborative inquiry learning. *Learning and Instruction, 17*(6), 708–721. doi:10.1016/j.learninstruc.2007.09.021

Lehman, M. E. (2007). Influence of learning style heterogeneity on cooperative learning. *NACTA Journal, 51*(4), 17–22.

Lehman, B., Matthews, M., D'Mello, S., & Person, N. (2008). What are you feeling? Investigating student affective states during expert human tutoring sessions. In B. P. Woolf, E. Aïmeur, R. Nkambou, & S. Lajoie (Eds.), *Intelligent Tutoring Systems* (pp. 50–59). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-69132-7_10

Linnenbrink, Elizabeth, A., & Pintrich, P. R. (2002). Achievement goal theory and affect: An asymmetrical bidirectional model. *Educational Psychologist, 37*(2), 69–78. doi:10.1207/S15326985EP3702_2

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Lund, K., Niccolai, G., Lavoué, E., Hmelo-Silver, C., Gweon, G., and Baker, M. (Eds.). (2019). *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019, Volume 1*. Lyon, France: International Society of the Learning Sciences. Retrieved from https://cscl2019.com/

Manske, S., Hecking, T., Chounta, I.-A., Werneburg, S., & Hoppe, H. U. (2015). Using differences to make a difference: A study on heterogeneity of learning groups. In O. Lindwall, P. Häkkinen, T. Koschmann, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 1* (pp. 182–189). Gothenburg, Sweden: The International Society of the Learning Sciences. Abstract retrieved from https://www.isls.org/cscl2015/

Marks, J., Copland, E., Loh, E., Sunstein, C. R., & Sharot, T. (2018). Epistemic spillovers: Learning others' political views reduces the ability to assess and use their expertise in nonpolitical domains. *Cognition*, 188, 74-84. doi:10.1016/j.cognition.2018.10.003

Marsh, H. W., & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology, 47*(1), 213–231. doi:10.1037/0022-3514.47.1.213

Marshall, P., Hornecker, E., Morris, R., Dalton, N. S., & Rogers, Y. (2008). When the fingers do the talking: A study of group participation with varying constraints to a tabletop interface. *3rd IEEE International Workshop on Horizontal Interactive Human Computer Systems, Amsterdam, Netherlands*, 33–40. doi:10.1109/TABLETOP.2008.4660181

Mercer, N. (1996). The quality of talk in children's collaborative activity in the classroom. *Learning and Instruction, 6*(4), 359–377. https://doi.org/10.1016/S0959-4752(96)00021-7

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Mercier, E. M., & Higgins, S. E. (2013). Collaborative learning with multi-touch technology: Developing adaptive expertise. *Learning and Instruction, 25*, 13–23. doi:10.1016/j.learninstruc.2012.10.004

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1999). Early fraction calculation ability. *Developmental Psychology, 35*(1), 164–174. doi:10.1037/0012-1649.35.1.164

Niedenthal, P. M., Barsalou, L. W., Winkielman, P., Krauth-Gruber, S., & Ric, F. (2005). Embodiment in attitudes, social perception, and emotion. *Personality and Social Psychology Review, 9*(3), 184–211. doi:10.1207/s15327957pspr0903_1.

Ottmar, E., & Landy, D. (2017). Concreteness fading of algebraic instruction: Effects on learning. *Journal of the Learning Sciences, 26*(1), 51–78. doi:10.1080/10508406.2016.1250212

Polo, C., Lund, K., Plantin, C., & Niccolai, G. P. (2016). Group emotions: The social and cognitive functions of emotions in argumentation. *International Journal of Computer-Supported Collaborative Learning, 11*(2), 123–156. doi:10.1007/s11412-016-9232-8

Puhl, T., Tsovaltzi, D., & Weinberger, A. (2015). A long-term view on learning to argue in Facebook: The effects of group awareness tools and argumentation scripts. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.). *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015: Volume 1* (pp. 110-117). Gothenburg, Sweden: The International Society of the Learning Sciences. Abstract retrieved from https://www.isls.org/cscl2015/

Randolph, J. J. (2008). *Multidisciplinary methods in educational technology research and development*. Hämeenlinna, Finland: HAMK University of Applied Sciences.

Rauscher, F. H., Krauss, R. M., & Chen, Y. (1996). Gesture, speech, and lexical access: The role of lexical movements in speech production. *Psychological Science, 7*(4), 226–231. doi:10.1111/j.1467-9280.1996.tb00364.x

Reinholz, D. L., Trninic, D., Howison, M., & Abrahamson, D. (2010). It's not easy being green: Embodied artifacts and the guided emergence of mathematical meaning. In P.

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Brosnan, D. B. Erchick, & L. Flevares (Eds.), *Proceedings of the 32nd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education: Optimizing Student Understanding in Mathematics* (pp. 1488–1496). Columbus, OH: The Ohio State University. Retrieved from https://www.pmena.org/

Reis, R. C. D., Isotani, S., Rodriguez, C. L., Lyra, K. T., Jaques, P. A., & Bittencourt, I. I. (2018). Affective states in computer-supported collaborative learning: Studying the past to drive the future. *Computers & Education, 120*, 29–50. doi:10.1016/j.compedu.2018.01.015

Resnick, L. B. (1989). Developing mathematical knowledge. *American Psychologist, 44*(2), 162–169. doi:10.1037/0003-066X.44.2.162

Rick, J. (2012, July). *Tiny tabletops: A research agenda*. Position paper presented at a workshop: Digital Ecosystems for Collaborative Learning (DECL 2012): Embedding Personal and Collaborative Devices to Support Classrooms of the Future, Sydney, Australia. Abstract retrieved from https://sites.google.com/site/iclsdecl/papers

Rick, J., Kopp, D., Schmitt, L., & Weinberger, A. (2015). Tarzan and Jane share an iPad. In O. Lindwall, P. Häkkinen, T. Koschman, P. Tchounikine, & S. Ludvigsen (Eds.), *Exploring the Material Conditions of Learning: The Computer Supported Collaborative Learning (CSCL) Conference 2015, Volume 1* (pp. 356–363). Gothenburg, Sweden: The International Society of the Learning Sciences. Abstract retrieved from https://www.isls.org/cscl2015/

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *Journal of the Learning Sciences, 2*(3), 235–276. doi:10.1207/s15327809jls0203_1

Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer Supported Collaborative Learning* (pp. 69–97). Berlin, Heidelberg, Germany: Springer. doi:10.1145/130893.952914

Rosebery, A. S., Ogonowski, M., DiSchino, M., & Warren, B. (2010). "The coat traps all your body heat": Heterogeneity as fundamental to learning. *Journal of the Learning Sciences, 19*(3), 322–357. doi:10.1080/10508406.2010.491752

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Schmitt, L. J., & Weinberger, A. (2018). Computer-supported collaborative learning: Mediated and co-present forms of learning together. In J. Voogt, G. Knezek, R. Christensen, & K.-W. Lai (Eds.), *Second handbook of information technology in primary and secondary education* (pp. 217–231). Cham: Springer. doi:doi.org/10.1007/978-3-319-53803-7_15-1

Schmitt, L. J., & Weinberger, A. (2019). Fourth graders' dyadic learning on multi-touch interfaces—versatile effects of verbalization prompts. *Educational Technology Research and Development*, *67*(3), 519–539. doi:10.1007/s11423-018-9619-5

Schneps, M. H., Ruel, J., Sonnert, G., Dussault, M., Griffin, M., & Sadler, P. M. (2014). Conceptualizing astronomical scale: Virtual simulations on handheld tablet computers reverse misconceptions. *Computers & Education, 70*, 269–280. doi:10.1016/j.compedu.2013.09.001

Schwarz, B. B., & Linchevski, L. (2007). The role of task design and argumentation in cognitive development during peer interaction: The case of proportional reasoning. *Learning and Instruction, 17*(5), 510–531. doi:10.1016/j.learninstruc.2007.09.009

Segal, A. (2011). *Do Gestural Interfaces Promote Thinking? Embodied Interaction: Congruent Gestures and Direct Touch Promote Performance in Math* (Doctoral dissertation). Retrieved from https://academiccommons.columbia.edu/catalog/ac:132260

Sim, G., MacFarlane, S., & Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. *Computers & Education, 46*(3), 235–248. doi:10.1016/j.compedu.2005.11.021

Spackman, J. S., & Yanchar, S. C. (2014). Embodied cognition, representationalism, and mechanism: A review and analysis. *Journal for the Theory of Social Behaviour, 44*(1), 46–79. doi:10.1111/jtsb.12028

Stahl, G. (2006). Supporting group cognition in an online math community: A cognitive tool for small-group referencing in text chat. *Journal of Educational Computing Research, 35*(2), 103–122. doi:10.2190/Q435-7611-2561-720P

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Strack, F., Martin, L. L., & Stepper, S. (1988). Inhibiting and facilitating conditions of the human smile: A nonobtrusive test of the facial feedback hypothesis. *Journal of Personality and Social Psychology, 54*(5), 768–777. doi:10.1037/0022-3514.54.5.768

ter Vrugte, J., de Jong, T., Vandercruysse, S., Wouters, P., van Oostendorp, H., & Elen, J. (2015). How competition and heterogeneous collaboration interact in prevocational game-based mathematics education. *Computers & Education, 89*, 42–52. doi:10.1016/j.compedu.2015.08.010

Tsovaltzi, D., Judele, R., Puhl, T., & Weinberger, A. (2015). Scripts, individual preparation and group awareness support in the service of learning in Facebook: How does CSCL compare to social networking sites? *Computers in Human Behavior*, *53*, 577–592. doi:10.1016/j.chb.2015.04.067

Tsovaltzi, D., Weinberger, A., Scheuer, O., Dragon, T., McLaren, B. M. (2013). Collaborative learning in Facebook: Can argument structure facilitate academic opinion change?. In N. Rummel, M. Kapur, M. Nathan, & S. Puntambekar (Eds.), *To See the World and a Grain of Sand: Learning Across Levels of Space, Time and Scale: CSCL 2013 Conference Proceedings Volume 2 – Short papers, Panels, Posters*, *Demos & Community Events* (pp. 177–180). Madison, WI: The International Society of the Learning Sciences. Abstract retrieved from https://www.isls.org/cscl/2013/

Tsovaltzi, D., Weinberger, A., Schmitt, L., Bellhäuser, H., Müller, A., Konert, J., … van Aalst, J. (2019). Group Formation in the Digital Age: Relevant Characteristics, their Diagnosis, and Combination for Productive Collaboration. In K. Lund, G. Niccolai, E. Lavoué, C. Hmelo-Silver, G. Gweon, & M. Baker (Eds.), *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning, Volume 2* (pp. 719-726). Lyon, France: The International Society of the Learning Sciences. Abstract retrieved from https://cscl2019.com/

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

van Dijk, A. M., Gijlers, H., & Weinberger, A. (2014). Scripted collaborative drawing in elementary science education. *Instructional Science*, *42*(3), 353–372. doi:10.1007/s11251-013-9286-1

Vogel, F., Reichersdorfer, E., Kollar, I., Ufer, S., Reiss, K., & Fischer, F. (2013). Learning to argue in mathematics: Effects of heuristic worked examples and CSCL scripts on transactive argumentation. In N. Rummel, M. Kapur, M. Nathan, & S. Puntambekar (Eds.), *To See the World and a Grain of Sand: Learning Across Levels of Space, Time and Scale: CSCL 2013 Conference Proceedings Volume 1 – Full Papers & Symposia* (pp. 526–533). Madison, WI: The International Society of the Learning Sciences. Abstract retrieved from https://www.isls.org/cscl/2013/

Webb, N. M. (1982). Group composition, group interaction, and achievement in cooperative small groups. *Journal of Educational Psychology, 74*(4), 475–484. doi:10.1037/0022-0663.74.4.475

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

Webb, N. M. (1993). Collaborative group versus individual assessment in mathematics: Processes and outcomes. *Educational Assessment, 1*(2), 131–152. doi:10.1207/s15326977ea0102_3

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

Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning and Instruction*, *17*(4), 416–426. doi:10.1016/j.learninstruc.2007.03.007

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*(4), 625–636. doi:10.3758/BF03196322

Schmitt, L. J., Tsovaltzi, D., & Weinberger A. (2019). *The role of heterogeneous cognitive and bodily processes – Learner interactions in embodied collaborative learning with tablets*. Manuscript in preparation.

Wosnitza, M., & Volet, S. (2012). Group heterogeneity and homogeneity in personal content goals for a group learning activity: Impact on individual appraisals. *Applied Psychology, 61*(4), 585–604. doi:10.1111/j.1464-0597.2012.