Computer-Supported Collaborative Learning

Sadhana Puntambekar
Gijsbert Erkens
Cindy Hmelo-Silver *Editors*

Analyzing Interactions in CSCL

Methods, Approaches and Issues

Springer

Analyzing Interactions in CSCL

COMPUTER-SUPPORTED COLLABORATIVE LEARNING

The *Computer-Supported Collaborative Learning Book Series* is for people working in the CSCL field. The scope of the series extends to 'collaborative learning' in its broadest sense; the term is used for situations ranging from two individuals performing a task together, during a short period of time, to groups of 200 students following the same course and interacting via electronic mail. This variety also concerns the computational tools used in learning: elaborated graphical whiteboards support peer interaction, while more rudimentary text-based discussion forums are used for large group interaction. The series will integrate issues related to CSCL such as collaborative problem solving, collaborative learning without computers, negotiation patterns outside collaborative tasks, and many other relevant topics. It will also cover computational issues such as models, algorithms or architectures which support innovative functions relevant to CSCL systems.

The edited volumes and monographs to be published in this series offer authors who have carried out interesting research work the opportunity to integrate various pieces of their recent work into a larger framework.

For other titles published in this series, go to
www.springer.com/series/5814

Sadhana Puntambekar • Gijsbert Erkens
Cindy E. Hmelo-Silver
Editors

Analyzing Interactions in CSCL

Methods, Approaches and Issues

Springer

Editors
Sadhana Puntambekar
University of Wisconsin
Madison, WI, USA
puntambekar@education.wisc.edu

Cindy E. Hmelo-Silver
Rutgers University
New Brunswick, NJ, USA
cindy.hmelo-silver@gse.rutgers.edu

Gijsbert Erkens
Research Centre Learning in Interaction
Utrecht University
Utrecht, The Netherlands
g.erkens@fss.uu.nl

ISBN 978-1-4419-7709-0 e-ISBN 978-1-4419-7710-6
DOI 10.1007/978-1-4419-7710-6
Springer New York Dordrecht Heidelberg London

Dedicated to our accomplished children
Aakash Puntambekar
Dr. Laura Hmelo
Dr. Samantha Hmelo

Acknowledgements

Like any edited volume, this book is a product of the combined efforts of several people. First, this book would not have been possible without the contribution of our colleagues who wrote the chapters. We learned from reading about the different perspectives the authors presented, and hope that our readers will do the same. We would especially like to thank our colleagues who wrote the commentaries for the book. Stephanie Teasley, Susanne Lajoie and Kristin Lund wrote the three commentaries, often at a short notice and when they had several other commitments. We thank them for their time and their valuable insights. Finally, we thank our students – Samaa Abdurraqib, Anushree Bopardikar and Dana Gnesdilow for their help with proofreading and formatting the chapters.

Introduction

Sadhana Puntambekar, Gijsbert Erkens, and Cindy E. Hmelo-Silver

Technology-enhanced environments to support student learning are becoming ubiquitous in both formal and informal educational contexts. Often, these environments require groups of students to learn collaboratively. As groups of learners engage in joint construction of meaning on a shared task, there is an emphasis on understanding how the group as a whole constructs knowledge through joint activity; this is a distinct shift from the traditional lens that views learning as a highly individualistic process and product. As Stahl et al. (2006) point out:

> Earlier studies of learning in groups treated learning as a fundamentally individual process. The fact that the individuals worked in groups was treated as a contextual variable that influenced the individual learning. In CSCL, by contrast, learning is also analyzed as a group process; analysis of learning at both the individual and the group unit of analysis are necessary. (p. 411)

This is what makes CSCL methodologically unique. This uniqueness is reflected in the several approaches that have been put forth to document and analyze collaborative interactions. CSCL as a field has made great strides from early research that focused on the extent of participation (De Wever et al. 2006). Currently, researchers use an array of qualitative and quantitative methods, including content analysis, social network analysis, analysis of log files, multilevel models, visual representations of data, etc., to analyze and model collaborative learning. Methods for analysis have included both an analysis of the process of learning and the learning outcomes. Further, measures of individual learning and learning by the group as a whole have been used.

This book is an attempt to discuss a representative set of current methods to analyze collaborative interactions, both at the individual and group levels. CSCL research tends to span across several disciplines such as education, psychology, computer science and artificial intelligence, bringing a diverse set of methods from research in these fields. The 15 chapters in this book present these diverse perspectives to provide researchers with a collection of methodologies to document and analyze collaborative interactions. A couple of recurring themes can be found through several of the chapters: unit of analysis, grain size of data, segmenting of data and temporality of interactions in CSCL. Additionally, several authors present frameworks that use multiple data sources and multiple methods of analysis.

One of the most important challenges of assessing collaborative learning is the issue of *unit of analysis*. Stahl et al. (2006) have pointed out that CSCL researchers are confronted with the issue of determining the appropriate unit for analysis. Establishing a unit of analysis poses difficulty because although group interactions are influenced by what the individual participants bring to the group, group processes are more than the sum of parts, and need to be understood as an entity within themselves. As Reimann (2007) described, learning in CSCL environments occurs "in individuals in the form of learning and in groups in the form of participation and knowledge building" (p. 611). Therefore, interactions can be analyzed with both the individual and the group as units of analysis. Further, within each level, the *grain size* of the unit needs to be determined based on the research questions that drive the analysis in a particular study (Chi 1997). Grain sizes can vary from analyzing a set of single utterances, chunks of discourse segmented along topics or themes, or both (e.g., Ash 2007; De Wever et al. 2006). These issues of units of analyses, grain size, and segmentation are addressed by several chapters in this volume.

Another recurring theme addressed in the book is that of *temporality* of data in CSCL environments. CSCL interactions occur over a period of time. Therefore analyzing single episodes does not adequately provide information about the process of learning. As Mercer (2008) described, "the coherence of educational experience is dependent on talk among participants, and so analyses of the ways that their continuing shared experience is represented and the ways that talk itself develops and coheres over an extended period are required" (p. 55).

The chapters are divided into three parts, as discussed in the next few paragraphs.

Part I: Understanding Group Processes

Kapur Voiklis, & Kinzer address the issue of uncovering temporal patterns in CSCL interactions by using a complex systems approach to the study of convergence in groups. In doing so, they address another significant aspect of group processes, that of divergence of ideas among group members and convergence of a group's understanding (Teasley et al. 2008; Roschelle 1992). Stahl (2004) argues that initial divergent ideas between group members significantly affect collaboration, because group members have to negotiate towards shared meaning. But it is not clear whether initially divergent groups eventually converge, and whether convergence is desirable. Nonetheless, despite the importance given to intersubjective meaning-making in the CSCL literature (Suthers 2006), convergence remains a difficult parameter to analyze. In their analysis, Kapur, Voiklis and Kinzer coded each discourse move in relation to the group goal, based on whether or not it moved the group toward a correct solution to the problem. A Markov model was then used to predict the group's performance.

Law, Yuen, Wong, and Leng discuss an approach to understand learners' trajectories in a group during asynchronous collaboration. They report a study in which

participation statistics were combined with specific aspects of discourse data in Knowledge Forum®. Law et al. propose a methodology to enable automatic coding and visualization of productive discourse threads on three interrelated aspects (scaffolds used, argumentative discourse markers and content topics).

Collaborative learning happens in a context, and several contextual variables affect a group's learning. Context can be broadly conceived as the physical and psychological variables that emerge from person-to-person interactions in any interpersonal human environment. Broad cultural influences of family and state intermix with more local cultures of schools and peer groups to provide additional constraints on how CSCL will emerge as a context for each individual learner. Arvaja addresses this issue by proposing a methodology based on sociocultural theories of learning, taking into account how physical and contextual aspects of any environment affect a group's discourse.

An important aspect of context is the tool that facilitates collaboration. Each tool has different affordances and has to be adapted to the context in which it is being used. Hmelo-Silver and Liu focus on the notion of how the effect of computer tools are important in mediating a group's discourse by visually representing data and taking into consideration the chronological relationship between talk and tool use. Their method uses a chronological representation of data, addressing the issue of temporality of CSCL interactions. Finally, Stahl argues for studying group cognition in CSCL and presents a case from the Math Forum project. As an example, he shows how proposals structure the temporal flow of the group interaction and thereby establish the social order of group cognition. Besides a temporal dimension, a problem space of shared knowledge artifacts and an interaction space of positioned actors are co-constructed by collaborative small-groups which define other dimensions of this social order. These group processes are, according to Stahl, not analyzable as individual behaviors, but can only be understood taking the group as unit of analysis.

Part II: Understanding Learning Within Groups

As mentioned earlier, CSCL interactions occur both at the level of the individual and groups. Contributions of individual group members influence themselves, other group members, and group processes as a whole. The ways in which individuals take up ideas and how the group as a whole moves forward are important aspects of CSCL to document and analyze, especially because each of these levels is unique for any group and also for the same group at different times. As such, it is important to study how membership in a group affects an individual member's learning, as well as the temporal aspects of how this learning changes over time. Both of these issues are addressed in the chapters in this part of the book.

Two of the chapters in this part focus on how the effect of individual membership in groups can be analyzed using multilevel models. Jansen, Erkens, Kirschner, and Kanselaar explain the use of multilevel modeling to account for both the individual and group level variables in the analysis of CSCL interactions. They discuss three

problems in analyzing data: hierarchically nested datasets, non-independence and differing units of analysis. They then illustrate strategies to address these problems through three examples. Stylianou-Georgiou, Papanastasiou, and Puntambekar further develop the idea of using multilevel modeling to analyze nested data by modeling the dependencies in their data to understand relationships between the variables of interest. In their study, they use both individual and group measures to apply a two-level model to understand the role of group membership in individual students' learning outcomes. Their analysis allows them to understand how attributes of the learning environment interact with group measures to affect individual learning outcomes.

Reimann, Yacef, and Kay address the issue of temporality by discussing how log data can be analyzed using data mining techniques. CSCL researchers often collect data in the form of log files to understand group interactions, resulting in large amounts of log data that need to be reduced, organized and analyzed. Log data often capture interactions that occur over time, such that events and sequences are related to each other. Reimann et al. address this issue by proposing data mining techniques that aim to identify sequence patterns and discrete event models. They propose analyzing group processes at an atomistic level as well as a holistic level. The chapter also acknowledges the challenges in applying data mining techniques, particularly the quality of data that is collected, and the level of granularity of that data, because "the quality of a model depends on the quality of the data" (Reimann et al. this volume). Thus, although data mining can be a powerful mechanism to analyze large quantities of log data, it is important to keep in mind the complexity of collecting the data in the first place and interpreting the resulting models.

The theme of temporal analysis continues in the chapter by Jeong, Clark, Sampson, and Menekse, who also propose the use of sequential analysis for group data. However, they use this approach with a coding of the discourse moves. Using a coding scheme for argumentative discourse moves, the sequential analysis helped them to identify, visualize, and assess the dialogic, temporal processes of argumentation in online science learning environments.

Part III: Frameworks for Analyzing Interaction in CSCL

The final part of this volume focuses on frameworks for analyzing collaborative interactions. Stegmann and Fischer present a model with heuristics for segmentation and coding. They discuss the challenging issue of segmenting, which is a key component of analyzing collaborative interactions. The grain size of the unit of analysis is an issue that all CSCL researchers grapple with, as we saw in earlier parts. Smaller segments in data provide finer grained analysis but little contextual information. On the other hand, larger units of analysis help create context but with the loss of detail. Therefore data segments need to be determined based on the research questions and goals for analysis (Chi 1997; Chavajay & Rogoff 2002).

Understanding group functioning as a whole, is a focus for several authors in this part of the book. This is a pertinent issue, because the quality of a group's interactions, often affect the outcomes. Groups that are dysfunctional may not accrue the benefits of learning collaboratively (e.g., Barron 2003). Rummel, Meier, Spada, Kahrimanis, and Avouris discuss the value of analyzing collaborative interactions as a whole, based on a few parameters that rate the quality of how well groups collaborate, such as communication, collaboration, etc. This is significant because failure of productive collaboration among group members can have a detrimental effect on both individual learning and collective knowledge development. Rummel et al. provide us with a tool that can be adapted and used by CSCL researchers, and can be flexibly combined with other measures, such as coding the utterances at a fine-grained level.

In a similar vein, Gweon, Jun, Lee, Finger, and Rosè tackle the difficult issue of the level and quality of interactions in face-to-face communication in groups. Difficulties in the interactional processes of groups often affect the outcomes of group work. The authors present their approach to tracking group progress in an instructional context. In a series of studies, they first identified process categories that needed to be tracked and then coded these during group work. The authors offer ways to automate this process using machine learning techniques. The approach that Gweon et al. discuss has implications for instructors and teachers to gain insight into group functions in classroom contexts. Suthers and Medina address a major topic of recent CSCL research, that of combining multiple logfiles of collaborative activities from different media and tools for analysis. They introduce the notion of contingency graphs that allow researchers to combine data that is distributed across media, enabling them to have a single abstract artifact with links to the original data.

Extending frameworks to the use of mobile devices, Scanlon describes case studies to showcase her approach to analyzing collaborative learning in several projects. Her framework, CIAO (Context, Interaction, Attitude and Outcomes) uses data collected from a variety of sources, both qualitative and quantitative. It is interesting to note that Scanlon discusses challenges to this approach as collaborative learning extends to mobile devices, creating a richer, social and technological setting. An Activity Theory framework seems to be promising to analyze the way activities are mediated by technology, as was illustrated in one of the presented studies. However, for the analysis of temporal aspects of knowledge construction a broader socio-cultural analytic approach is suggested by the author. Finally, Martínez-Monés, Harrer, Dimitriadis report on the requirements for computers to support CSCL researchers in conducting interaction analysis. The chapter notes limitations in trying to conduct analysis post-hoc in existing tools. The authors propose a design process for the development of CSCL environments. For example, a *co-design* approach in which learning and analysis needs are integrated from the start or a *multi-perspective* approach in which these two needs are treated independently at an initial stage and integrated later can be employed. Technology requirements and solutions to support this integration of learning and analysis are discussed as well.

Conclusion

Together, the chapters in this book provide a suite of tools that can be applied, modified and customized to document and analyze collaborative interactions. There are of course issues that still need to be explored. For example, while we have a range of methods for assessing learning outcomes and group processes, the issue of measuring group outcomes as a whole still remains a challenge (see Lund this volume). Our hope is to start the conversation on the different methods discussed, as the CSCL community moves forward to find the best ways to understand individual learning and group processes in collaborative environments.

References

Ash, D. (2007). Using video data to capture discontinuous science meaning making in non-school settings. In R. Goldman, R. Pea, B. Barron, & S. Derry (Eds.), *Video research in the learning sciences*. Mahwah: Erlbaum.

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, *12*(3), 307–359.

Chavajay, P., & Rogoff, B. (2002). Schooling and traditional collaborative social organization of problem solving by Mayan mothers and children. *Developmental Psychology*, *38*(1), 55–66.

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

De Wever, B., Schellens T., Valcke, M., & Van Keer H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education*, *46*(1), 6–28.

Mercer, N. (2008). The seeds of time: why classroom dialogue needs a temporal analysis. *Journal of the Learning Sciences*, *17*(1), 33–59.

Mercer, N., Littleton, K., & Wegerif, R. (2004) Methods for studying the processes of interaction and collaborative activity in computer-based educational activities. *Technology, Pedagogy and Education*, *13*(2), 193–209.

Reimann, P. (2007). Time is precious: Why process analysis is essential for CSCL (and also can help to bridge between experimental and descriptive methods). In C. Chinn, G. Erkens, & S. Puntambekar (Eds.), *Minds, minds, and society. Proceedings of the Computer-supported Collaborative Learning Conference (CSCL 2007)* (pp. 598–607). New Brunswick: International Society of the Learning Sciences.

Roschelle, J. (1992). Learning by collaborating: Convergent conceptual change. *The Journal of the Learning Sciences*, *2*(3), 235–276.

Stahl, G. (2004) Building collaborative knowing- elements of a social theory of learning. In J. W. Strijbos, P. A. Kirschner, & R. L. Martens (Eds.), *What we know about CSCL in higher education* (pp. 53–86) Dordrecht: Kluwer.

Stahl, G., Koschmann, T., & Suthers, D. D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 409–426). Cambridge: Cambridge University Press.

Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning (ijCSCL)*, *1*(3), 315–337.

Teasley, S. D., Fischer, F., Weinberger, A., Stegmann, K., Dillenbourg, P., Kapur, M., et al. (2008). Cognitive convergence in collaborative learning. In *Proceedings of the 8th international conference on the learning sciences* (Vol. 3, pp. 360–367). Mahwah: Lawrence Erlbaum.