Computer-Supported Collaborative Learning (2009) 4:115–149
DOI 10.1007/s11412-009-9061-0

The joint organization of interaction within a multimodal CSCL medium

Murat Perit Çakır · Alan Zemel · Gerry Stahl

Received: 14 June 2008 / Accepted: 19 February 2009 / Published online: 26 March 2009

Abstract In order to collaborate effectively in group discourse on a topic like mathematical patterns, group participants must organize their activities in ways that share the significance of their utterances, inscriptions, and behaviors. Here, we report the results of a ethnomethodological case study of collaborative math problem-solving activities mediated by a synchronous multimodal online environment. We investigate the moment-by-moment details of the interaction practices through which participants organize their chat utterances and whiteboard actions as a coherent whole. This approach to analysis foregrounds the sequentiality of action and the implicit referencing of meaning making—fundamental features of interaction. In particular, we observe that the sequential construction of shared drawings and the deictic references that link chat messages to features of those drawings and to prior chat content are instrumental in the achievement of intersubjectivity among group members' understandings. We characterize this precondition of collaboration as the co-construction of an indexical field that functions as a common ground for group cognition. Our analysis reveals methods by which the group co-constructs meaningful inscriptions in the dual-interaction spaces of its CSCL environment. The integration of graphical, narrative, and symbolic semiotic modalities in this manner also facilitates joint problem solving. It allows group members to invoke and operate with multiple realizations of their mathematical artifacts, a characteristic of deep learning of mathematics.

Keywords Group cognition · Interaction analysis · Dual-interaction space · Ethnomethodology · Indexicality · Mathematics education · Text chat · Visual reasoning · Common ground · Joint problem space

M. Perit Çakır (✉) · G. Stahl
College of Information Science & Technology, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA
e-mail: mpc48@drexel.edu

G. Stahl
e-mail: gerry.stahl@ischool.drexel.edu

A. Zemel
The Department of Culture & Communication, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA
e-mail: arz26@drexel.edu

Computer-supported collaborative learning is centrally concerned with the joint organization of interaction by small groups of students in online environments. The term "collaborative learning" is a gloss for *interaction that is organized for the joint achievement of knowledge-building tasks* such as problem solving in domains like school mathematics. Rather than using the term "collaborative learning," which carries vague and contradictory connotations, we coined the term "group cognition" to refer to activities where several students organize their joint interaction to achieve such collective cognitive accomplishments as planning, deducing, designing, describing, problem solving, explaining, defining, generalizing, representing, remembering, and reflecting as a group.

We have argued in *Group Cognition* (Stahl 2006) that CSCL interactions should be analyzed at the group level of description, not just at the individual or the community levels, as is done in other theoretical approaches influential in CSCL research. During the past six years, we have conducted the Virtual Math Teams (VMT) Project to explore group cognition in a prototypical CSCL setting and to analyze it at the group level. We have used our analyses of interaction to drive the design of the technology.

In this paper, we present a case study of an 18-minute-long excerpt from the VMT Project. We look at some ways in which the students organized their joint efforts. Our observations here are consistent with our impressions from more than a hundred student-hours of interaction in the VMT data corpus. Many of the broader theoretical and practical issues surrounding the analysis here are addressed by CSCL researchers in a new edited volume on *Studying Virtual Math Teams* (Stahl 2009b) in the Springer CSCL book series.

The issue that we address in the following pages is: *How do the students in our case study organize their activity so they can define and accomplish their tasks as a group within their online environment?* This is necessarily a pivotal question for a science of CSCL (Stahl 2009a). It involves issues of meaning making, shared understanding and common ground that have long been controversial in CSCL.

The problem of coordination is particularly salient in the VMT software environment, which is an instance of a dual-interaction space (Dillenbourg 2005; Mühlpfordt and Stahl 2007), requiring organization across multiple media, each with their own affordances. We have found that the key to joint coordination of knowledge building is sequential organization of a network of indexical and semantic references within the group discourse (Stahl 2007). We therefore analyze sequential interaction at the group level of description, using ethnomethodologically inspired chat interaction analysis rather than quantitative coding, in order to maintain and study this sequential organization. Thereby, we arrive at a view of mathematical knowledge building as the coordinated production and use of visual, narrative, and symbolic inscriptions as multiple realizations of co-constructed mathematical objects.

While we have elsewhere presented theoretical motivations for focusing on *group discourse organization* as fundamental for CSCL, in this paper we foreground our *analysis of empirical data* from a VMT session. We derive a number of characteristics of the joint organization of interaction from the details of the case study. The characteristics we describe are to some extent specific to the technological affordances of the VMT environment, to the pedagogical framing of the chat session, and even to the unique trajectory of this particular group interaction. Nevertheless, the characteristics are indicative of what takes place—with variations—in similar settings. After the analytic centerpiece of the paper, we discuss *methodological implications* for CSCL analysis, including what it means to take the *group* as the unit of analysis. We then contrast our approach to leading *alternative approaches* in CSCL. This discussion focuses particularly on multimodal interaction in a *dual-interaction space* and on related conceptions of *common ground*, concluding with summary remarks on *sequential analysis*. The paper proceeds through the following topics:

- The problem of group organization in CSCL
- A case study of a virtual math team
- Implications for CSCL chat interaction analysis
- The group as the unit of analysis
- Other approaches in CSCL to analyzing multimodal interaction
- Grounding through interactional organization
- Sequential analysis of the joint organization of interaction

The problem of group organization in CSCL

A central issue in the theory of collaborative learning is how students can solve problems, build knowledge, accomplish educational tasks, and achieve other cognitive accomplishments *together*. How do they share ideas and talk about the same things? How do they know that they are talking about, thinking about, understanding, and working on things in the same way? Within CSCL, this has been referred to as the problem of the "attempt to construct and maintain a shared conception of a problem" (Roschelle and Teasley 1995), "building common ground" (Baker et al. 1999; Clark and Brennan 1991) or "the practices of meaning making" (Koschmann 2002). We have been interested in this issue for some time. *Group Cognition* (Stahl 2006) documents a decade of background to the VMT research reported here: Its Chapter 10 (written in 2001) argued the need for a new approach and its Chapter 17 (written in 2002) proposed the current VMT Project, which includes this case study. Since 2002, we have been collecting and analyzing data on how groups of students in a synchronous collaborative online environment organize their interaction to achieve intersubjectivity and shared cognitive accomplishments in the domain of school mathematics.

Knowledge building in CSCL has traditionally been supported primarily with asynchronous technologies (Scardamalia and Bereiter 1996). Within appropriate educational cultures, this can be effective for long-term refinement of ideas by learning communities. However, in small groups and in many classrooms, asynchronous media encourage mere exchange of individual opinions more than co-construction of progressive trains of joint thought. We have found informally that synchronous interaction can more effectively promote group cognition—the accomplishment of "higher order" cognitive tasks through the coordination of contributions by individuals within the discourse of a small group. We believe that the case study in this paper demonstrates the power of group interaction in a largely synchronous environment; the coordination of interaction in an asynchronous interaction would be quite different in nature as a result of very different interactional constraints.

In CSCL settings, interaction is mediated by a computer environment. Students working in such a setting must enact, adapt, or invent ways of coordinating their understandings by means of the technological affordances that they find at hand (see Dohn, this issue). The development and deployment of these methods is not usually an explicit, rational process that is easily articulated by either the participants or analysts. It occurs tacitly, unnoticed, taken-for-granted. In order to make it more visible to us as analysts, we have developed an environment that makes the coordination of interaction more salient and captures a complete record of the group interaction for detailed analysis. In trying to support online math problem solving by small groups, we have found it important to provide media for both linguistic and graphical expression. This resulted in what is known within CSCL as a *dual-interaction space*. In our environment, students must coordinate their text chat postings with their whiteboard drawings. A careful analysis of how they do this reveals as well their more general methods of group organization.

The analysis of our case study focuses on episodes of interaction through which an online group of students co-constructs mathematical artifacts across dual-interaction spaces. It looks closely at how group members put the multiple modalities into use, how they make their chat postings and drawing actions intelligible to each other, and how they achieve a sense of coherence among actions taking place across the modalities to which they have access. We base our discussion, analysis, and design of the affordances of the online environment on the methodical ways the features of the software are put into use by the students.

In another VMT case study (Sarmiento and Stahl 2008), we have seen how the problem-solving work of a virtual math team is accomplished through the co-construction and maintenance of a *joint problem space* (Teasley and Roschelle 1993). This figurative space—that supports group interaction and the shared understanding of that interaction by the participants—not only grounds the *content* of the team's discourse and work, but also ties together the *social* fabric of the relations among the team members as actors. In addition, we saw that the joint problem space has a third essential dimension: *time* or sequence. The construction of the joint problem space constitutes a shared temporality through bridging moves that span and thereby order discontinuous events as past, present, and future (Sarmiento-Klapper 2009). This can be seen, for instance, in the use of tenses in group-remembering discourses. More generally, the joint problem space provides a framework of sequential orderings, within which temporal deictic references, for example, can be resolved.

In this paper, we further investigate how a virtual math team achieves a group organization of its activities such that the group can proceed with a sense of everyone understanding each other and of working collaboratively as a group. We do this through a fine-grained analysis of the group's interaction in a VMT session in which they formulate, explore, and solve a geometry problem. Their work takes place in graphical, narrative, and symbolic media—supported technologically by the shared whiteboard, text chat, and wiki pages of the VMT environment. We pay particular attention to how graphical inscriptions, textual postings, and symbolic expressions in the different media are closely coordinated by the group members, despite the differences of the media.

We pursue a micro-ethnographic approach to analyzing the activities of the group members in their own terms. They set themselves a task, propose how to proceed step by step, and explain to each other how to understand their actions. We try to follow the explanations, which are available in the inscriptions, postings, and expressions—particularly when the sequentiality of these allows the complex references among them to be followed.

The establishment of group order in small-group interaction is always strongly dependent upon the media, which mediate interaction. In the case of VMT chats, there is an intricate set of technological media, including text chat, a shared whiteboard, a community wiki, and graphical references from chat to whiteboard. The central part of this paper explores the different characteristics of the VMT media by observing how the students use them. Of particular interest are the ways in which a group coordinates activities in the different graphical and textual media. From a math-education perspective, it is also insightful to see how the visual and narrative understandings feed into the development and understanding of symbolic expressions.

By the end of the paper, we will see how the group organization of graphical, narrative, and symbolic resources in interaction continuously produce and reproduce the joint problem space of the group's effort. This coordination is revealed through sequential analysis, in which the consequence of one action in one medium following another in another medium is seen as mutually constitutive of the meaning of those actions. The sequential web of activity across the VMT media—woven by semantic and indexical references among them—forms the joint problem space within which problem content, participant relationships, and temporal progress are all defined in a way that is shared by the group. We can see the "indexical field" (Hanks 1992) formed by the group activities as the source of grounding that supports the intersubjectivity of the group effort. In contrast to psychological or psycholinguistic models of common ground, the fact that team members believe they have understandings in common about what each other is saying and doing is not a result of exchanging individual mental opinions, but is a function of the indexical organization of the group interaction.

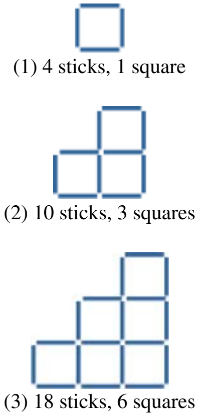
The joint problem space—as the foundation of group cognition—is not a mental construct of a set of individuals who achieve cognitive convergence or common (identical) ground through comparing mental models anymore than it is a figment of some form of group mind. Rather, it is a system of interconnected meanings formed by a weaving of references in the group discourse itself (Stahl 2007). In this paper, we analyze the methods the students used to co-construct this indexical field.

In our case study, the organization of group meaning making takes place across media—in accordance with the specific affordances of the different media. Furthermore, the grounding of the students' symbolic mathematical understanding can be seen as related to their visual and narrative understandings—or, rather, the various understandings are intricately interwoven and support each other. We trace this interweaving through our approach to the interactional analysis of sequential coordination at the group unit of analysis.

A case study of a virtual math team

The excerpts we present in this paper are obtained from a problem-solving session of a team of three students who participated in the VMT Spring Fest 2006. This event brought together several teams from the US, Scotland, and Singapore to collaborate on an open-ended math task on geometric patterns. Students were recruited anonymously through their teachers. Members of the teams generally did not know each other before the first session. Neither they nor we knew anything about each other (e.g., age or gender) except chat handle and information that may have been communicated during the sessions. Each group participated in four sessions during a two-week period, and each session lasted over an hour. An adult from the research project moderated each session; the facilitators' task was to help the teams when they experienced technical difficulties, not to participate in the problem-solving work.

During their first session, all the teams were asked to work online on a particular pattern of squares made up of sticks (see Fig. 1). For the remaining three sessions the teams were asked to come up with their own shapes, describe the patterns they observed as mathematical formulas, and share their observations with other teams through a wiki page. This task was chosen because of the possibilities it afforded for many different solution approaches ranging from simple counting procedures to more advanced methods involving the use of recursive functions and exploring the properties of various number sequences. Moreover, the task had both algebraic and geometric aspects, to allow us to observe how participants put many features of the VMT software system into use. The open-ended nature of the activity stemmed from the need to agree upon a new shape made by sticks. This required groups to engage in an open-ended problem-solving activity, as compared to traditional situations where questions are given in advance and there is a single "correct" answer—presumably already known by a teacher. We used a traditional pattern problem (Moss and Beatty 2006; Watson and Mason 2005) to seed the activity and then left it up to each group to decide the kinds of shapes they found interesting and worth exploring further.



N	Sticks	Squares
1	4	1
2	10	3
3	18	6
4	?	?
5	?	?
6	?	?
...	...	...
N	?	?

Session I

1. Draw the pattern for N=4, N=5, and N=6 in the whiteboard. Discuss as a group: How does the graphic pattern grow?
2. Fill in the cells of the table for sticks and squares in rows N=4, N=5, and N=6. Once you agree on these results, post them on the VMT Wiki
3. Can your group see a pattern of growth for the number of sticks and squares? When you are ready, post your ideas about the pattern of growth on the VMT Wiki.

Sessions II and III

1. Discuss the feedback that you received about your previous session.
2. **WHAT IF?** Mathematicians do not just solve other people's problems - they also explore little worlds of patterns that they define and find interesting. Think about other mathematical problems related to the problem with the sticks. For instance, consider other arrangements of squares in addition to the triangle arrangement (diamond, cross, etc.). **What if** instead of squares you use other polygons like triangles, hexagons, etc.? Which polygons work well for building patterns like this? How about 3-D figures, like cubes with edges, sides and cubes? What are the different methods (induction, series, recursion, graphing, tables, etc.) you can use to analyze these different patterns?
3. Go to the VMT Wiki and share the most interesting math problems that your group chose to work on.

Fig. 1 Task description

All the problem-solving sessions were conducted in the VMT environment. The VMT online system has two main interactive components that conform to the typical layout of systems with dual-interaction spaces: a shared drawing board that provides basic drawing features on the left, and a chat window on the right (Fig. 2). The online environment has features specifically designed to help users relate the actions happening across dual-interaction spaces (Stahl 2009b, chap.15). One of the unique features of this chat system is the referencing support mechanism (Mühlpfordt and Wessner 2005) that allows users to visually connect their chat postings to previous postings or objects on the whiteboard via arrows (see the last posting in Fig. 2 for an example of a message-to-whiteboard reference). The referential links attached to a message are displayed until a new message is posted. Messages with referential links are indicated by an arrow icon in the chat window, and a user can see where such a message is pointing by clicking on it at any time.

In addition to the explicit referencing feature, the system displays small boxes in the chat window to indicate actions performed on the whiteboard. This awareness mechanism allows users to observe how actions performed in both interaction spaces are sequenced with respect to each other. Moreover, users can click on these boxes to move the whiteboard back and forth from its current state to the specific point in its history when that action was performed. Chat messages and activity markers are color coded to help users to keep track of who is doing what

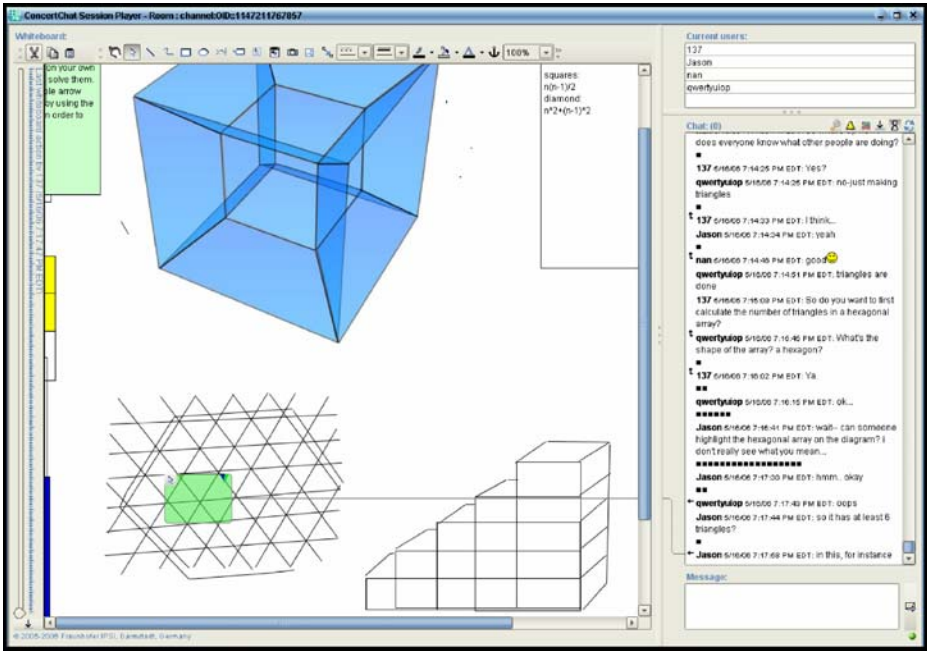


Fig. 2 A screen-shot of the VMT environment

in the online environment. In addition to standard awareness markers that display who is present in the room and who is currently typing, the system also displays textual descriptions of whiteboard actions in tool-tip messages that can be observed by holding the mouse either on the object in the whiteboard or on the corresponding square in the chat window.

Studying the meaning-making practices enacted by the users of CSCL systems inevitably requires a close analysis of the process of collaboration itself (Dillenbourg et al. 1996; Stahl et al. 2006). In an effort to investigate the organization of interactions across the dual-interaction spaces of the VMT environment, we consider the small group as the unit of analysis (Stahl 2006), and we appropriate methods of ethnomethodology and conversation analysis to conduct sequential analysis of group interactions at a microlevel (Psathas 1995; Sacks 1962/1995; ten Have 1999). Our work is informed by studies of interaction mediated by online text chat with similar methods (Garcia and Jacobs 1998, 1999; O'Neill and Martin 2003), although the availability of a shared drawing area and explicit support for deictic references in our online environment substantially differentiate our study from theirs.

The goal of this line of analytic work is to discover the commonsense understandings and procedures group members use to organize their conduct in particular interactional settings (Coulon 1995). Commonsense understandings and procedures are subjected to analytical scrutiny because they are what "enable actors to recognize and act on their real world circumstances, grasp the intentions and motivations of others, and achieve mutual understandings" (Goodwin and Heritage 1990, p. 285). Group members' shared competencies in organizing their conduct not only allow them to produce their own actions, but also to interpret the actions of others (Garfinkel and Sacks 1970). Because group members enact these understandings visibly in their situated actions, researchers can discover them through detailed analysis of the members' sequentially organized conduct (Schegloff and Sacks 1973).

We conducted numerous VMT Project data sessions, where we subjected our analysis of the excerpts below to intersubjective agreement (Psathas 1995). This paper presents the outcome of this group effort together with the actual transcripts so that the analysis can be subjected to external scrutiny. During the data sessions we used the VMT Replayer tool, which allows us to replay a VMT chat session as it unfolded in real time based on the time stamps of actions recorded in the log file. The order of actions—chat postings, whiteboard actions, awareness messages—we observe with the Replayer as researchers exactly matches the order of actions originally observed by the users. This property of the Replayer allowed us to study the sequential unfolding of events during the entire chat session, which is crucial in making sense of the complex interactions mediated by a CSCL environment (Koschmann et al. 2007).

In this case study, we focus on a sequence of excerpts obtained from a single problem-solving session of a virtual math team. We are concerned with how the actors contribute to the group meaning making as they proceed. This example involves the use and coordination of actions involving both the whiteboard and chat environment. It therefore served as a useful site for seeing how actors, in this local setting, were able to engage in meaningful coordinated interaction.

The team has three members: Jason, 137 and Qwertyuiop, who are upper-middle-school students (roughly 14 years old) in the US. In the following subsections, we will present how this team co-constructed a mathematical artifact they referred to as the “hexagonal array” through a coordinated sequence of actions distributed between the chat and whiteboard spaces, and how they subsequently explored its properties by referring to and annotating shared drawings on the whiteboard. In particular, we will highlight how whiteboard objects and previous chat postings were used as semiotic resources during the collaborative problem-solving activity. This will show how chat and whiteboard differ in terms of their affordances for supporting group interaction. We will see how these differences are enacted and used in complementary ways by team members to achieve mutual intelligibility of their actions across multiple interaction spaces.

Availability of production processes

Log 1 is taken from the beginning of the team’s third session. The team has already explored similar patterns of sticks and become familiar with the features of the VMT online environment during their prior sessions. The drawing actions at the beginning of this excerpt were the first moves of the session related to math problem solving.

Log 1

Line	Time	Chat handle	Chat message or <*whiteboard action*>
	7:07:52–7:11:00	137	<*137 draws a hexagon shape and then splits it up into regions by adding lines. Figure 3 shows some of the key steps in 137’s drawing performance*>
1	7:11:16	137	Great. Can anyone m ake a diagram of a bunch of triangles?
	7:11:16–7:11:49	137	<*137 deletes the set of lines he has just drawn*>
2	7:11:51	Qwertyuiop	just a grid?….
	7:11:54–7:12:01	137	<*137 moves some of the older drawings away*>
3	7:12:07	137	Yeah...
4	7:12:17	Qwertyuiop	ok…
	7:12:23–7:14:07	Qwertyuiop	<*Qwertyuiop draws a grid of triangles in the space opened up by 137. Figure 4 shows some of the steps in Qwertyuiop’s drawing actions*>

At the beginning of this excerpt, 137 performs a series of drawing actions. 137's actions on the whiteboard include the drawing of a hexagon first, then three diagonal lines and finally lines parallel to the diagonals and to the sides of the hexagon whose intersections eventually introduce some triangular and diamond-shaped regions. Moreover, 137 also performs some adjustment moves—for instance between the 4th and 5th snapshots in Fig. 3—to ensure that three non-parallel lines intersect at a single point, and the edges of the hexagon are parallel to the lines introduced later as much as possible. Hence, this sequence of drawing actions suggests a particular organization of lines for constructing a hexagonal shape. (Fig. 3 shows six snapshots corresponding to intermediary stages of 137's drawing actions: 137 initiates his drawing actions with six lines that form the hexagon in stage 1. Then he adds three diagonal lines in step 2. The 3rd snapshot shows the additional two lines drawn parallel to one of the diagonals. The 4th snapshot shows a similar set of two parallel lines added with respect to another diagonal. The 5th snapshot shows slight modifications performed on the new set of parallel lines to ensure intersections at certain places. The 6th snapshot shows the final stage of 137's drawing.)

137's chat posting in line 1 that follows his drawing effort (which can be read as a self-critical, sarcastic "great") suggests that he considers his illustration inadequate in some way. He makes this explicit by soliciting help from other members to produce "a diagram of a bunch of triangles" on the whiteboard, and then removing the diagram he has just produced (the boxes following this posting in Fig. 5 correspond to deletion actions on the whiteboard). By removing his diagram, 137 makes that space available to other members for the projected drawing activity. Qwertyuiop responds to 137's query with a request for clarification regarding the projected organization of the drawing ("just a grid?"). After 137's acknowledgement, Qwertyuiop performs a series of drawing actions that resemble the latter stages of 137's drawing actions, namely starting with the parallel lines tipped to the right

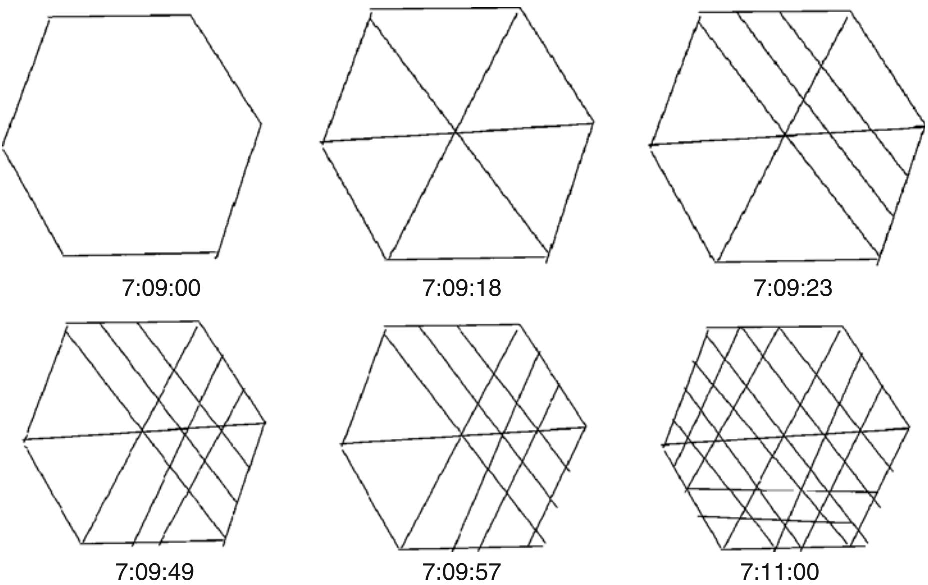


Fig. 3 Six stages of 137's drawing actions obtained from the Replayer tool. The time stamp of each stage is displayed under the corresponding image. Snapshots focus on a particular region on the whiteboard where the relevant drawing activity is taking place

first, then drawing a few parallel lines tipped to the left, and finally adding horizontal lines at the intersection points of earlier lines that are parallel to each other (see Figs. 4 and 5). Having witnessed 137's earlier actions, the similarity in the organizations of both drawing actions suggest that Qwertyuiop has appropriated some key aspects of 137's drawing strategy, but modified/reordered the steps (e.g., he did not start with the hexagon at the beginning) in a way that allowed him to produce a grid of triangles as a response to 137's request.

The key point we would like to highlight in this episode is that *the availability of the sequencing of the drawing actions that produces a diagram on the shared whiteboard can serve as a vital resource for collaborative sense-making*. As seen in Log 1, 137 did not provide any explanation in chat about his drawing actions or about the shape he was trying

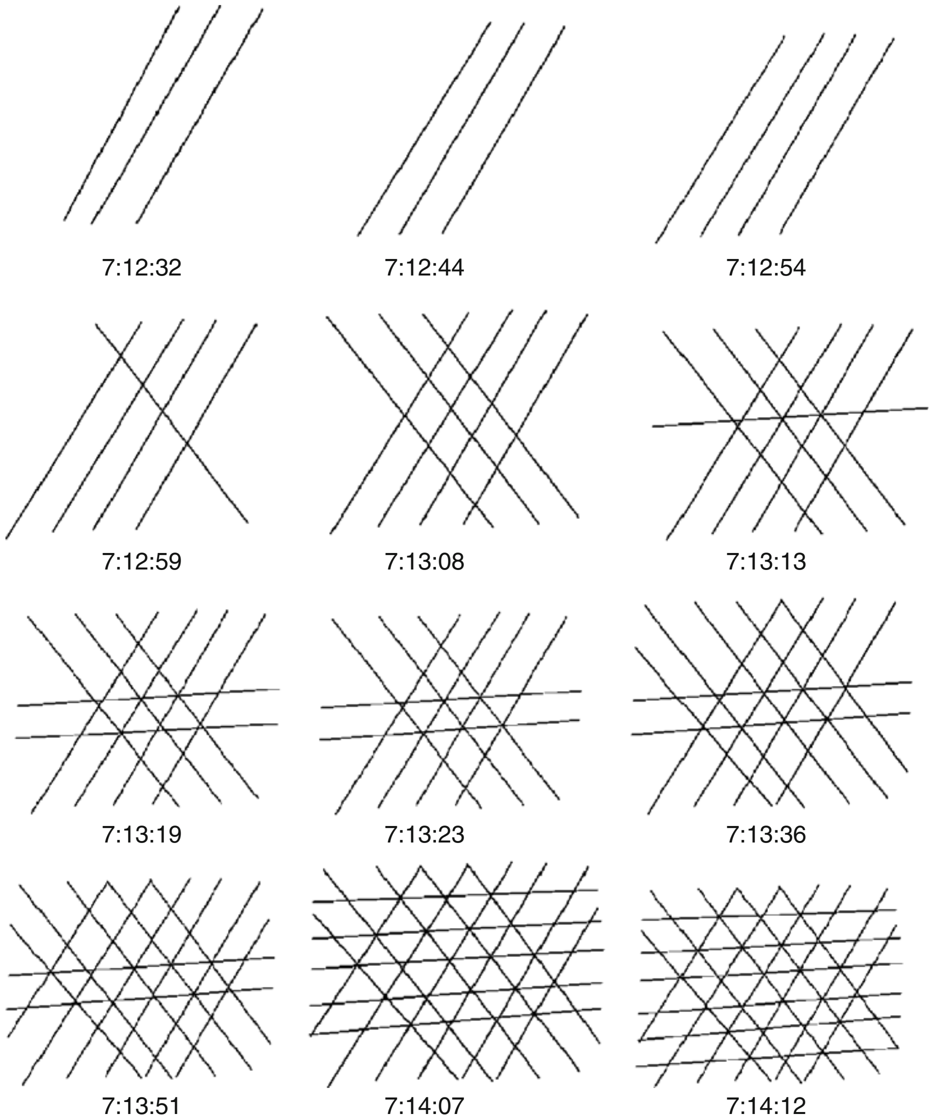


Fig. 4 The evolution of Qwertyuiop's drawing in response to 137's request

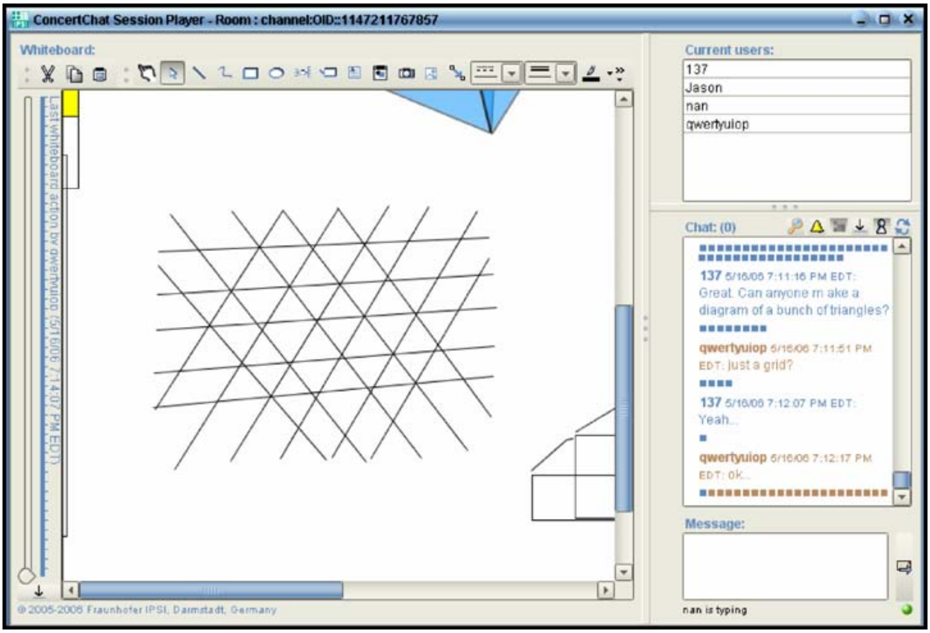


Fig. 5 The interface at the 12th stage of Fig. 4

to draw. Yet, as we have observed in the similarity of Figs. 3 and 4, the orderliness of 137's actions has informed Qwertyuiop's subsequent performance. The methodical use of intersecting parallel lines to produce triangular objects is common to both drawing performances. Moreover, Qwertyuiop does not repeat the same set of drawing actions, but selectively uses 137's steps to produce the relevant object (i.e., a grid of triangles) on the whiteboard. Qwertyuiop does not initially constrain his representational development by constructing a hexagon first, but allows a hexagon (or other shapes made with triangles) to emerge from the collection of shapes implied by the intersecting lines. Thus, Qwertyuiop's performance shows us that he is able to *notice a particular organization* in 137's drawing actions, and he has *selectively appropriated and built upon* some key aspects of 137's drawing practice. As we will see in the following logs,[1] the group's subsequent use of this drawing will provide us additional evidence that Qwertyuiop's diagram serves as an adequate response to 137's request.

This excerpt highlights a fundamental difference between the two interaction spaces: whiteboard and chat contributions differ in terms of the availability of their production process. As far as chat messages are concerned, participants can only see who is currently typing,[2] but not what is being typed until the author decides to send the message. A similar

[1] For instance, after Qwertyuiop declares the completion of the grid in line 11, 137 anchors Qwertyuiop's drawing to the background at 7:15:47 (see Log 3). Because such a move preserves the positions of the selected objects and the objects affected by the move include only the lines recently added by Qwertyuiop, 137's anchoring move seems to give a particular significance to Qwertyuiop's recent drawing. Hence, 137's anchoring move can be treated as an (implicit) endorsement of Qwertyuiop's drawing effort in response to his previous request.

[2] While a participant is typing, a social awareness message appears under the chat entry box on everyone else's screen stating that the person "is typing" (see Fig. 5). When the typist posts the message, the entire message appears suddenly as an atomic action in everyone's chat window.

situation applies to *atomic* whiteboard actions such as drawing an individual line or a rectangle. Such actions make a single object appear in the shared drawing area when the user releases the left mouse button; in the case of editable objects such as textboxes, the object appears on the screens of the computers of all chat participants when the editor clicks outside the textbox. However, the construction of most shared diagrams includes the production of multiple atomic shapes (e.g., many lines), and hence the sequencing of actions that produce these diagrams is available to other members. As we have observed in this excerpt, the availability of the drawing process can have interactionally significant consequences for math-problem-solving chats due to its instructionally informative nature. In short, the whiteboard affords an *animated evolution* of the shared space, which makes the *visual reasoning process* manifest in drawing actions *publicly available* for other members' inspection. For instance, in Fig. 4, transitions from stages 1 to 2 and 7 to 8 show modifications performed to achieve a peculiar geometric organization on the shared workspace.

Mutability of chat and whiteboard contents

Another interactionally significant difference between the chat and the whiteboard interaction spaces, which is evidenced in the excerpt above, is the difference in terms of the mutability of their contents. Once a chat posting is contributed, it cannot be changed or edited. Moreover, the sequential position of a chat posting cannot be altered later on. If the content or the sequential placement of a chat posting turns out to be interactionally problematic, then a new posting needs to be composed to repair that. On the other hand, the object-oriented design of the whiteboard allows users to reorganize its content by adding new objects and by moving, annotating, deleting, and reproducing existing ones. For instance, the way 137 and Qwertyuiop repaired their drawings in the excerpt above by repositioning some of the lines they drew earlier to make sure that they intersect at certain points and/or that they are parallel to the edges of the hexagon illustrates this difference. Such demonstrable tweaks make the mathematical details of the construction work visible and relevant to observers, and hence, serve as a vital resource for joint mathematical sense making. By seeing that Qwertyuiop successively and intentionally adjusts lines in his whiteboard drawing to appear more parallel or to intersect more precisely, the other group members take note of the significance of the arrangement of lines as parallel and intersecting in specific patterns.

While both chat and whiteboard in VMT support persistence, visibility, and mutability, they do so in different ways. A chat posting scrolls away only slowly and one can always scroll back to it, whereas a drawing may be erased by anyone at any time. Chat conventions allow one to replace (i.e., follow) a mistyped posting with a new one, and conversational conventions allow utterances to be retracted, repaired, or refined. The mechanisms of the two mediational technologies are different and the characteristics of their persistence, visibility, and mutability differ accordingly. Collaborative interaction in the dual-space environment is sensitively attuned to these intricate and subtle differences.

Monitoring joint attention

The excerpt in Log 2 immediately follows the one in Log 1, where the team is oriented to the construction of a triangular grid after a failed attempt to embed a grid of triangles inside a hexagon. As Qwertyuiop is adding more lines to the grid, the facilitator (Nan) posts two questions addressed to the whole team in line 5. The question not only queries about what is happening now and whether everybody knows what others are currently doing, but the placement of the question at this point in interaction also problematizes the relevance of what has been happening so far. 137's response in lines 6 and 8 treat the facilitator's question as a problematic intervention. Qwertyuiop's response indicates he is busy with making triangles, and hence may not know what others are doing. Jason acknowledges that he is following what has been going on in line 9. These responses indicate that the team members have been following (perhaps better than the facilitator) what has been happening on the whiteboard so far as something relevant to their task at hand.

Log 2

5	7:14:09	nan	so what's up now? does everyone know what other people are doing?
	7:14:12	Qwertyuiop	< *Qwertyuiop adds a line to the grid of triangles*>
6	7:14:25	137	Yes?
7	7:14:25	Qwertyuiop	no-just making triangles
	7:14:32	Qwertyuiop	< *Qwertyuiop adds a line to the grid of triangles*>
8	7:14:33	137	I think... [REF to line 6]
9	7:14:34	Jason	Yeah
	7:14:36	Qwertyuiop	< *Qwertyuiop adds a line to the grid of triangles*>
10	7:14:46	nan	good :-)
11	7:14:51	Qwertyuiop	Triangles are done
12	7:15:08	137	So do you want to first calculate the number of triangles in a hexagonal array?

In this excerpt, the facilitator calls on each participant to report on his/her understanding of the activities of other participants. There was an extended duration in which no chat postings were published while whiteboard actions were being performed by Qwertyuiop. Because it is not possible for any participant to observe other participants directly, it is not possible to monitor a class of actions others may perform that (1) are important for how we understand ongoing action but (2) do not involve explicit manipulation of the VMT environment, actions like watching the screen, reading text, inspecting whiteboard constructs, and so forth. The only way to determine if those kinds of actions are occurring is to explicitly inquire about them using a chat posting.

Past and future relevancies implied by shared drawings

Following Qwertyuiop's announcement in line 11 of Log 2 that the drawing work is complete, 137 proposes that the team calculate "the number of triangles" in a "hexagonal array" as a possible question to be pursued next. Although a hexagon was previously produced as part of the failed drawing, this is the first time someone explicitly mentions the term "hexagonal array" in this session. What makes 137's proposal potentially intelligible to others is the availability of referable resources such as whiteboard objects, and the immediate history of the production of those objects such that the proposal can be seen to be embedded in a sequence of displayed actions. 137's use of "So" to

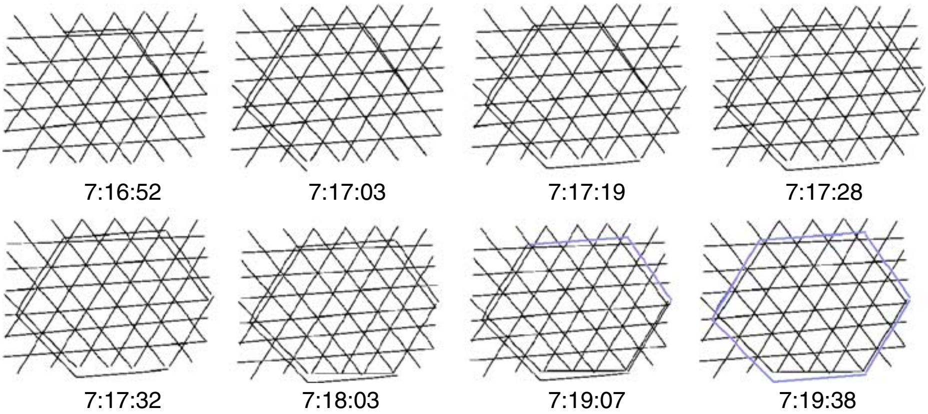


Fig. 6 Snapshots from the sequence of drawing actions performed by 137

introduce his proposal presents it as a consequence of, or a making explicit of, what preceded. His suggestion of it as a "first" (next) move implies that the drawings opened up multiple mathematical tasks that the group could pursue, and that the proposed suggestion would be a candidate for a next move. In other words, the objects on the whiteboard and their visually shared production index a horizon of past and future activities. The indexical terms in 137's proposal (like "hexagonal array") not only rely on the availability of the whiteboard objects to propose a relevant activity to pursue next, but also modify their sense by using linguistic and semantic resources in the production to label or gloss the whiteboard object and its production. This allows actors to orient in particular ways to the whiteboard object and the procedures of its co-construction—providing a basis for coordinated joint activity. The joint activity acquires a temporal structure that is defined by the details of chat wording, the animation of graphical construction, and the sequentiality of proposing.

Methods for referencing relevant objects in the shared visual field

Bringing relevant mathematical objects to other members' attention often requires a coordinated sequence of actions performed in both the chat and whiteboard interaction spaces. The episode following 137's proposal (Log 3) provides us with an appropriate setting to illustrate how participants achieve this in interaction. Following 137's proposal in line 12, both Qwertyuiop and Jason post queries for clarification in lines 13 and 16, respectively, which indicate that the available referential resources were insufficient for them to locate what 137 is referring to with the term "hexagonal array." Jason's query in the chat is particularly important here because it explicitly calls for a response to be performed on the shared diagram, that is, in a particular field of relevance in the other interaction space. Following Jason's query, 137 begins to perform a sequence of drawing actions on the shared diagram. He adds a few lines that gradually begin to enclose a region on the triangular grid[3] (see Fig. 6).

[3] In the meantime, Qwertyuiop also performs a few drawing actions near the shared drawing, but his actions do not introduce anything noticeably different because he quickly erases what he draws each time.

Log 3

11	7:14:51	Qwertyuiop	Triangles are done
12	7:15:08	137	So do you want to first calculate the number of triangles in a hexagonal array?
13	7:15:45	Qwertyuiop	What’s the shape of the array? a hexagon? *<REF to 12>*
	7:15:47	137	*<137 locks the triangular grid that Qwertyuiop has just drawn>*
14	7:16:02	137	Ya *<REF to line 13>*
15	7:16:15	Qwertyuiop	ok….
	7:16:18–7:16:35	137	*<137 performs a few drawing actions and then erases them>*
16	7:16:41	Jason	wait– can someone highlight the hexagonal array on the diagram? i don’t really see what you mean...
	7:16:45–7:17:28	137	*<137 adds new lines to the grid on the whiteboard which gradually forms a contour on top of the grid. Figure 6 shows some of the steps performed by 137>*
17	7:17:30	Jason	Hmm.. okay
18	7:17:43	Qwertyuiop	Oops *<REF to Whiteboard>*
19	7:17:44	Jason	so it has at least 6 triangles?
20	7:17:58	Jason	in this, for instance *<REF to Whiteboard>*
	7:18:03–7:18:17	137	*<137 completes the contour by adding more lines, which forms a hexagon>*
21	7:18:53	137	How do you color lines?
22	7:19:06	Jason	There’s a little paintbrush icon up at the top
23	7:19:12	Jason	it’s the fifth one from the right
	7:19:13–7:19:20	137	*137 begins to change the color of the lines that form the contour to blue>*
24	7:19:20	137	Thanks.
25	7:19:21	Jason	There ya go :-)
	7:19:25–7:19:48	137	*<137 finishes the coloring. Now the contour is highlighted in blue>*
26	7:19:48	137	Er... That hexagon.
27	7:20:02	Jason	so... should we try to find a formula i guess

When the shared diagram reaches the stage illustrated by the 4th frame in Fig. 6, Jason posts the message “hmmm… okay” in line 17, which can be read as an acknowledgement of 137’s performance on the whiteboard as a response to his recent chat query. Because no chat message was posted after Jason’s request in line 16, and the only shared actions were 137’s work on the whiteboard, Jason’s chat posting can be read as a response to the ongoing drawing activity on the whiteboard. As it is made evident in his posting, Jason is treating the evolving drawing on the shared diagram as a response to his earlier query for highlighting the hexagonal array on the whiteboard: The question/answer adjacency pair is spread across the two interaction spaces in an unproblematic way.

Following provisional acknowledgement of 137’s drawing actions on the whiteboard, Jason posts a claim in line 19. This posting is built as a declarative: “so it has at least 6 triangles,” with a question mark appended to the end. The use of “so” in this posting invites readers to treat what follows in the posting as a consequence of the prior actions of 137. In this way, Jason is (a) proposing a defeasible extension of his understanding of the sense of 137's actions and (b) inviting others to endorse or correct this provisional claim about the hexagonal array by presenting this as a query using the question mark.

In line 20, Jason provides further specificity to what he is indexing with the term "it" in line 19 by highlighting a region on the grid with the referencing tool of the VMT system. The textual part of the posting makes it evident that the highlighted region is an instance of the object mentioned in line 19. Moreover, the six triangles highlighted by the explicit reference recognizably make up a hexagon shape altogether. Hence, Jason's explicit reference seems to be pointing to a particular stage (indexed by "at least") of the hexagonal array to which the team is oriented (see Fig. 7).

In other words, having witnessed the production of the hexagonal shape on the whiteboard as a response to his earlier query, Jason displays his competence by demonstrating his recognition of the hexagonal pattern implicated in 137's graphical illustration. 137's drawing actions highlight a particular stage of a growing pattern made of triangles—stage $N=3$, as we will see in Fig. 9. However, recognizing the stick-pattern implicated in 137's highlighting actions requires other members to project how the displayed example can be grown and/or shrunk to produce other stages of the hexagonal array. Thus, Jason's description of the shape of the "hexagonal array" at a different stage—$N=1$—is a public display of his newly achieved comprehension of the significance of the math object in the whiteboard and the achievement of "indexical symmetry" among the parties involved with respect to this math object (see Stahl 2009b, chap.14).

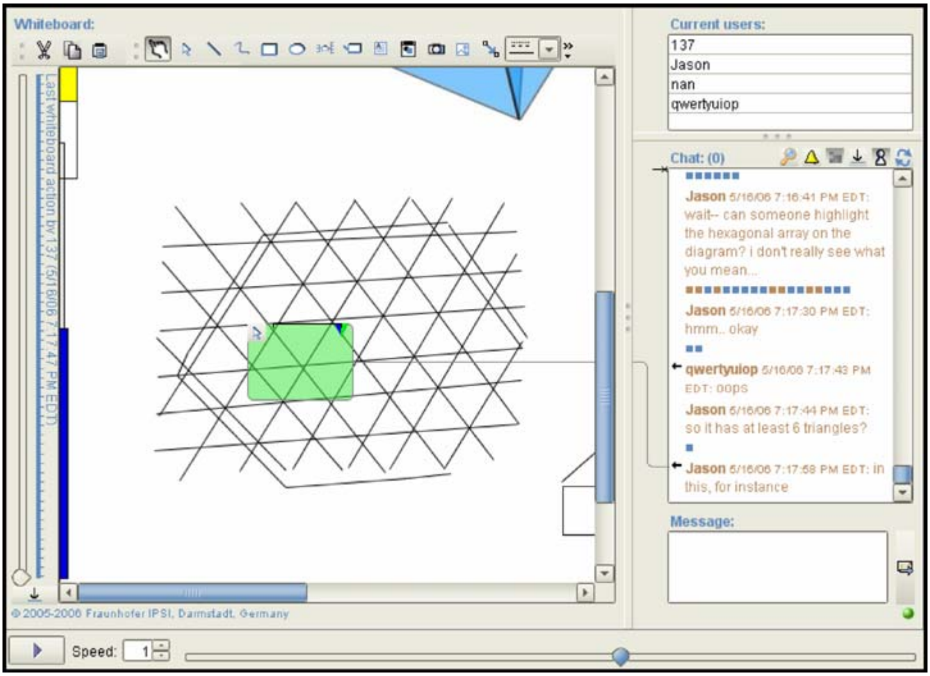


Fig. 7 Use of the referencing tool to point to a stage of the hexagonal array

Although Jason explicitly endorsed 137’s drawing as an adequate illustration, the small boxes in the chat stream that appear after Jason’s acknowledgement in line 17 show that 137 is still oriented to and operating on the whiteboard. In line 21, 137 solicits other members’ help regarding how he can change the color of an object on the board, which opens a side sequence about a specific feature of the whiteboard system. Based on the description he got, 137 finishes marking the hexagon by coloring all its edges with blue, and he posts “that hexagon” in line 25. This can be read as a chat reference to the whiteboard shape enclosed by the blue contour, and as a response to other members’ earlier requests for clarification.

In this excerpt, we have observed two referential methods enacted by participants to bring relevant graphical objects on the whiteboard to other group members’ attention. In the first case, 137 *marked the drawing* with a different color to identify the contour of a hexagonal shape. As evidenced in other members’ responses, this was designed to make the hexagonal array embedded in a grid of triangles visible to others. Jason demonstrated another method by using the explicit referencing tool to support his *textual description* of the first stage of the pattern. Both mechanisms play a key role in directing other members’ attention to features of the shared *visual field* in particular ways. This kind of deictic usage isolates components of the shared drawing and constitutes them as relevant objects to be attended to for the purposes at hand. As we shall see, these guided shifts in visual focus of the group have strategic importance for the group’s mathematical work. Hence, such referential work establishes a fundamental *relationship between the narrative and mathematical terminology used in text chat and the animated graphical constructions produced on the whiteboard*. The shared sense of the textual terms and the inscriptions co-evolve through the referential linkages established as the interaction sequentially unfolds in both interaction spaces.

In Log 3, the group tentatively proposes a major mathematical insight—that a hexagon can be viewed as six symmetric triangular areas. It is a visual achievement. It emerges from a visual inspection by Jason of 137’s graphical diagram, based on Qwertyuiop’s method of visually representing hexagons as patterns of triangularly intersecting lines. By literally focusing his eyes on a smallest hexagon in the larger array and counting the number of triangles visible within a hexagonal border, Jason discovers that there are at least six triangles at the initial stage of a hexagon with one unit on each side. We will see how the group visualizes the generalization of this picture to other stages. However, it is already interesting to note that Jason not only observes the composition of a small hexagon out of six triangles, but he conveys this to the rest of the group in both media. He posts chat line 19 and then references from chat line 20 to a visually highlighted view in the whiteboard, so that his *visual* understanding can be shared by the group as well as his *narrative* description in his claim. The next step for the group will be to formulate a *symbolic* mathematical expression of this claim.

Whiteboard visualizations, chat narratives and wiki symbolisms

The excerpt in Log 4 immediately follows Log 3. The way 137 uses both interaction spaces in this episode highlights another important aspect of collaborative problem-solving work in an environment like VMT. Because participants can contribute to only one of the interaction spaces at a time, they cannot narrate their whiteboard actions simultaneously with chat postings, as can be done with talking about a whiteboard in a face-to-face setting. However, as we will observe in 137’s use of the whiteboard in the following excerpt, participants can achieve a similar interactional organization by coordinating their actions in such a way that whiteboard actions can be seen as part of an exposition performed in chat.

Log 4

27	7:20:02	Jason	so... should we try to find a formula i guess
28	7:20:22	Jason	input: side length; output: # triangles
29	7:20:39	Qwertyuiop	It might be easier to see it as the 6 smaller triangles.
30	7:20:48	137	Like this? <*REF to line 29*>
	7:20:53	137	<*137 draws a red line*>
	7:20:57	137	<*137 draws a red line*>
	7:21:00	137	<*137 draws a red line*>
31	7:21:02	Qwertyuiop	Yes
32	7:21:03	Jason	Yup
	7:21:03	137	<*137 moves the second red line*>
	7:21:05	137	<*137 moves the second red line again. It is positioned on the grid now*>
33	7:21:29	Qwertyuiop	Side length is the same...
34	7:22:06	Jason	Yeah

Jason brings the prior activity of locating the hexagonal array on the shared drawing to a close with his so-prefaced posting in line 27, where he invokes the task of finding a formula that was mentioned by 137 earlier. Jason provides further specificity to the formula he is referring to in the next line (i.e., given the side length as input the formula should return the number of triangles as output). In line 29, Qwertyuiop takes up Jason's proposal by suggesting the team consider the hexagonal array as six smaller triangles to potentially simplify the task at hand. In the next line, 137 posts a question phrased as "like this?" which is addressed to Qwertyuiop's prior posting, as indicated by the use of the referential arrow. Next, we observe the appearance of three red lines on the shared diagram, which are all added by 137. Here, 137 demonstrates a particular way of splitting the hexagon into six

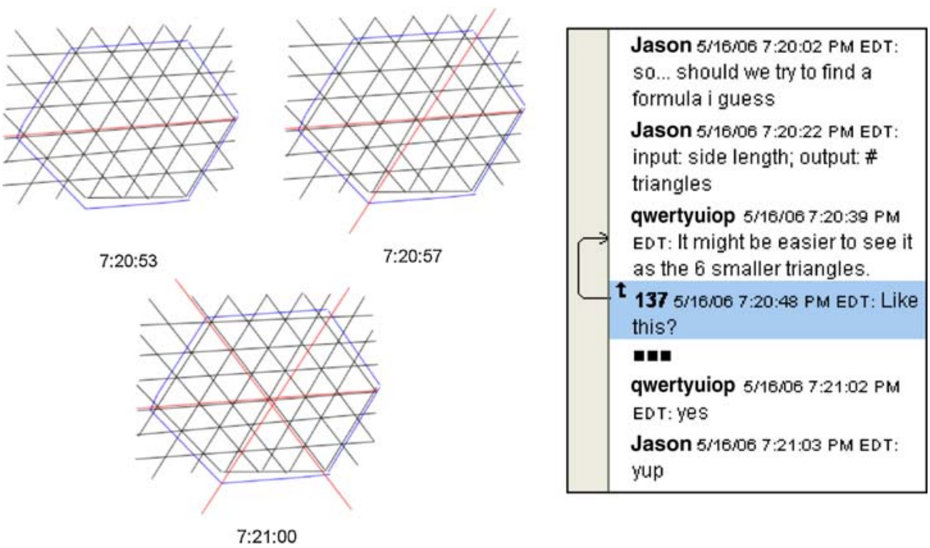


Fig. 8 137 splits the hexagon into six parts

parts: The image on the left of Fig. 8 corresponds to the sequence of three whiteboard actions represented as three boxes in the chat excerpt. After 137 adds the third line whose intersection with the previously drawn red lines recognizably produces six triangular regions on the shared representation, Qwertyuiop and Jason both endorse 137's demonstration of a particular way of splitting up the hexagonal shape.

One important aspect of this organization is directing other members' attention to the projected whiteboard activity as a relevant step in the sequentially unfolding exposition in chat. For instance, the deictic term "this" in 137's chat line 30 refers to something yet to be produced, and thereby projects that there is more to follow the current posting, possibly in the other interaction space. Moreover, the use of the referential link and the term "like" together inform others that what is about to be done should be read in relation to the message to which 137 is responding. Finally, 137's use of a different color marks the newly added lines as recognizably distinct from what is already there as the background, and hence, noticeable as a demonstration of what is implicated in recent chat postings.

Again, the progress in understanding the mathematics of the problem is propelled through visual means. In response to Jason's proposal of finding a formula, Qwertyuiop suggests that "it might be easier to see it" in a certain way. Jason's proposed approach might be difficult to pursue because no one has suggested a concrete approach to constructing a formula that would meet the general criteria of producing an output result for any input variable value. By contrast, the group has been working successfully in the visual medium of the whiteboard drawing and has been literally able to "see" important characteristics of the math object that they have co-constructed out of intersecting lines. Jason has pointed out that at least six triangles are involved (in the smallest hexagon). So, Qwertyuiop proposes building on this in-sight. 137 asks if the way to see the general case in terms of the six small triangles as proposed by Qwertyuiop can be visualized by intersecting the hexagon array with three intersecting lines to distinguish the six regions of the array. He does this through a visual construction, simply referenced from the chat with his "Like this?" post.

By staring at the final version of the array (stage 3 in Fig. 8), all members of the group can see the hexagon divided into six equal parts at each stage of the hexagonal pattern. Near the intersection of the red lines, they can see a single small triangle nestled in each of the six regions. As will be evidenced in Log 5, within the larger hexagon delimited by the blue lines, they can see a set of 1+3+5=9 small triangles in each of the six larger triangular regions. Similarly, midway between stage $N=1$ and stage $N=3$, one can visually observe 1+3=4 small triangles in each region. The new view, scaffolded by 137's red lines, entails *visual reasoning* that leads to mathematical deductions. As soon as Qwertyuiop and Jason see 137's construction, they both concur with it as the easier way to see the mathematical pattern of triangles in the hexagonal array. The visual reasoning supported by whiteboard and narrated textually in the chat will lead in the next episode to symbolic reasoning for posting in the wiki.

A first glance at the chat logs might suggest that the group is narrating their problem-solving process in the chat and illustrating what they mean by "napkin" drawings in the whiteboard, to use Dillenbourg and Traum's (2006) metaphor. However, a second look reveals that the most significant insight and sharing is occurring in the whiteboard, more along the lines of a visual "model" metaphor. Perhaps the best way to describe what is going on is to say that the group is very carefully coordinating their work in the dual space as a whole to achieve a shared progression of understanding of the pattern problem. This is accomplished with an efficiency and effectiveness that could not be achieved in either a purely textual chat system or a purely graphical whiteboard. Although in this view the chat and whiteboard both function as symmetric parts of a coordinated whole in which chat references drawing and drawing illustrates chat, it is important to differentiate their roles as well.

Using representations of specific instances as a resource for generalization

Immediately following the previous excerpt, the team moves on to figuring out a general formula to compute the number of triangles in a hexagonal pattern. In line 34 of Log 5, Jason relates the particular partitioning of the hexagon illustrated on the whiteboard to the problem at hand by stating that the number ("#") of triangles in the hexagon will equal 6 times ("×6") the number of triangles enclosed in each partition. In the next posting, 137 seems to be indexing one of the six partitions with the phrase "each one." Hence, this posting can be read as a proposal about the number of triangles included in a partition. The sequence of numbers in the expression "1+3+5" calls others to look at a partition in a particular way. While 137 could have simply said here that there are nine triangles in each partition, he instead organizes the numbers in summation form and offers more than an aggregated result. His expression also demonstrates a systematic method for counting the triangles. In other words, his construction is designed to highlight a particular orderliness in the organization of triangles that form a partition. Moreover, the sequence includes increasing consecutive odd numbers, which implicitly informs a certain progression for the growth of the shape under consideration.

Log 5

34	7:22:13	Jason	so it'll just be ×6 for # triangles in the hexagon
35	7:22:19	137	Each one has 1+3+5 triangles.
36	7:22:23	Jason	but then we're assuming just regular hexagons
37	7:22:29	Qwertyuiop	the "each polygon corrisponds to 2 sides" thing we did last time doesn't work for triangles
38	7:23:17	137	It equals 1+3+...+(n+n−1) because of the "rows"?
39	7:24:00	Qwertyuiop	yes- 1st row is 1, 2nd row is 3...
40	7:24:49	137	And there are n terms so... n(2n/2)
41	7:25:07	137	or n^2 <*REF to line 40*>
42	7:25:17	Jason	Yeah
43	7:25:21	Jason	then multiply by 6
44	7:25:31	137	To get 6n^2 <*REF to line 43*>

About a minute after his most recent posting, 137 offers an extended version of his sequence as a query in line 38. The relationship between the sequence for the special case and this one is made explicit through the repetition of the first two terms. In the new version the "…" notation is used to substitute a series of numbers following the second term up to a generic value represented by "n+n−1," which can be recognized as a standard expression for the n^{th} odd number. Hence, this representation is designed to stand for something more general than the one derived from the specific instance illustrated on the whiteboard. 137 attributes this generalization to the concept of "rows," and solicits other members' assessment regarding the validity of his version (by ending with a question mark). 137's use of the term "rows" seems to serve as a pedagogic device that attempts to locate the numbers in the sequence on the n^{th} stage of the hexagonal pattern (see Fig. 9 for an analyst's illustration of the generalized hexagonal pattern). For stages 1, 2, and 3, the hexagonal shape has 6*(1) = 6, 6*(1+3) = 24, 6*(1+3+5) = 54 triangles, respectively.

Qwertyuiop's endorsement of 137's proposal comes in line 39. He also demonstrates a row-by-row iteration on a hexagon, where each number in the sequence corresponds to a

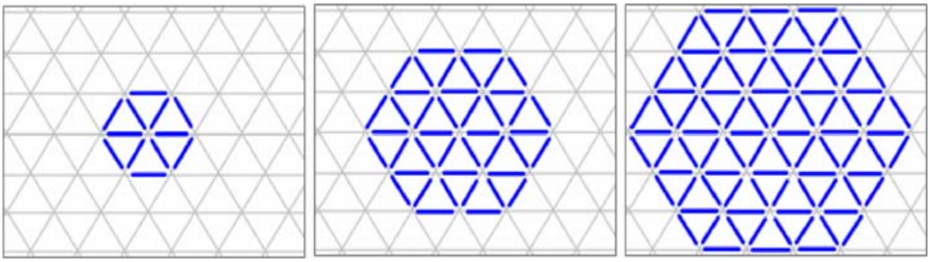


Fig. 9 A reconstruction of the first three iterations of the geometric pattern

row of triangles in a partition. In other words, Qwertyuiop elaborates on 137's statement in line 38 of the chat by displaying his understanding of the relationship between the rows and the sequence of odd numbers. Although he does not explicitly reference it here, Qwertyuiop may be viewing the figure in the whiteboard to see the successive rows. The figure is, of course, also available to 137 and Jason to help them follow Qwertyuiop's chat posting and check it.

Then 137 proposes an expression for the sum of the first n odd numbers in line 40.[4] Jason agrees with the proposed expression and suggests that it should be multiplied by 6 next. In the following line, 137 grammatically completes Jason's posting with the resulting expression. In short, by virtue of the agreements and the co-construction work of Jason and 137, the team demonstrates its endorsement of the conclusion that the number of triangles would equal $6n^2$ for a hexagonal array made of triangles. As the group collaboratively discovered, when n equals the stage number (as "input" to the formula), the number of triangles is given by the expression $6n^2$.

The way team members orient themselves to the shared drawing in this episode illustrates that the drawings on the whiteboard have a figurative role in addition to their concrete appearance as illustrations of specific cases. The particular cases captured by concrete, tangible marks on the whiteboard are often used as a resource to investigate and talk about general properties of the mathematical objects indexed by them.

Another important aspect of the team's achievement of a general expression in this episode is the way they transformed a particular way of *counting* the triangles in one of the partitions (i.e., a geometric observation) into an algebraic mode of investigation. This shift from a visual method led the team members to recognize that a particular sequence of numbers can be associated with the way the partition grows in subsequent iterations. The shift to this symbolic mode of engagement, which heavily uses the shared drawing as a resource, allowed the team to go further in the task of generalizing the pattern of growth by invoking algebraic resources. In other words, the team made use of multiple realizations (graphical and linguistic) of the math object (the hexagonal array) distributed across the dual-interaction space to co-construct a general formula for the task at hand.

Chat versus whiteboard contributions as persistent referential resources

In all of the excerpts we have considered so far, the shared drawing has been used as a resource within a sequence of related but recognizably distinct activities. For instance, the

[4] 137 makes use of Gauss's method for summing this kind of series, adding the first and last term and multiplying by half of the number of terms: $(1 + n + n - 1)*n/2 = 2n*n/2 = n^2$. This method was used by the group and shared in previous sessions involving the stair pattern that is still visible in the whiteboard.

group has oriented itself to the following activities: (1) drawing a grid of triangles, (2) formulating a problem that relates a hexagonal array to a grid of triangles, (3) highlighting a particular hexagon on the grid, (4) illustrating a particular way to split the shape into six smaller pieces, and (5) devising a systematic method to count the number of triangles within one of the six pieces. As the group oriented to different aspects of their shared task, the shared diagram was modified on the whiteboard and annotated in chat accordingly. Yet, although it had been modified and annotated along the way, the availability of this shared drawing on the screen and the way participants organize their discussion around it highlights its persistent characteristic as an ongoing referential resource. In contrast, none of the chat postings in prior excerpts were attributed a similar referential status by the participants. As we have seen, in each episode the postings responded or referred either to recently posted chat messages or to the visual objects in the shared space.

The textual chat postings and the graphical objects produced on the whiteboard differ in terms of the way they are used as referential resources by the participants. The content of the whiteboard is persistently available for reference and manipulation, whereas the chat content is visually available for reference for a relatively shorter period. This is due to the linear growth of chat content, which replaces previous messages with the most recent contributions inserted at the bottom of the chat window. Although one can make explicit references to older postings by using the scroll-bar feature, the limited size of the chat window affords a referential locality between postings that are visually (and hence temporally) close to each other.

By contrast, objects drawn in the whiteboard tend to remain there for a long time. They are often only erased or moved out of view when space is needed for drawings related to a new topic. While they may be modified, elaborated, or moved around, whiteboard objects may remain visible for an entire hour-long session or even across sessions. Like the chat, the whiteboard has a history scrollbar, so that any past state of the drawing can be made visible again—although in practice students rarely use this feature. Although both media technically offer a persistent record of their contents, the visual locality of the whiteboard—the fact that graphical objects tend to stay available for reference from the more fleeting chat—qualifies it as the more persistent medium as an interactional resource. This notion of persistence does not imply that the shared sense of whiteboard objects is fixed once they are registered to the shared visual field. As they continue to serve as referential resources during the course of the problem-solving effort, the sense of whiteboard objects may become increasingly evident and shared, or their role may be modified as participants make use of them for varying purposes.

Implications for CSCL chat interaction analysis

In this case study, we investigated how a group of three upper-middle-school students put the features of an online environment with dual-interaction spaces into use as they collaboratively worked on a math problem they themselves came up with. Our analysis has revealed important insights regarding the affordances of systems with dual-interaction spaces. First, we observed that the whiteboard can make visible to everyone the animated evolution of a geometric construction, displaying the *visual reasoning* process manifested in drawing actions. Second, whiteboard and chat contents differ in terms of *mutability* of their contents, due to the object-oriented design of the whiteboard that allows modification and annotation of past contributions. Third, the media differ in terms of the *persistence* of their contents: Whiteboard objects remain in the shared visual field until they are removed, whereas chat content gradually scrolls off as new postings are produced. Although contents of both spaces are persistently available for reference, due to linear progression of the chat window, chat postings are likely to refer to visually (and hence temporally) close chat messages and to graphical whiteboard objects. Finally, the whiteboard objects *index* a horizon of past and future activities as they serve as an interactional resource through the course of recognizably distinct but related episodes of chat discussion.

Our analysis of this team's joint work has also revealed methods for the organization of collaborative work, through which group members co-construct mathematical meaning sedimented in semiotic objects distributed across the dual-interaction spaces of the VMT environment. We observed that bringing relevant math artifacts referenced by indexical terms such as "hexagonal array" to other members' attention often requires a coordinated sequence of actions across the two interaction spaces. Participants use explicit and verbal references to guide each other about how a new contribution should be read in relation to prior contents. Indexical terms stated in chat referring to the visible production of shared objects are instrumental in the reification of those terms as meaningful mathematical objects for the participants. Verbal references to co-constructed objects are often used as a resource to index complicated and abstract mathematical concepts in the process of co-constructing new ones. Finally, different representational affordances of the dual-interaction spaces allow groups to develop multiple realizations of the math artifacts to which they are oriented. Shared graphical inscriptions and chat postings are used together as semiotic resources in mutually elaborating ways. Methods of coordinating group interaction across the media spaces also interrelate the mathematical significances of the multiple realizations.

Overall, we observed that actions performed in both interaction spaces constitute an evolving historical context for the joint work of the group. What gets done now informs the relevant actions to be performed next, and the significance of what was done previously can be modified depending on the circumstances of the ongoing activity. As the interaction unfolds sequentially, the sense of previously posted whiteboard objects and chat statements may become evident and/or refined. In this way, the group's joint problem space is maintained.

Through the sequential coordination of chat postings and whiteboard inscriptions, the group successfully solved their mathematical challenge, to find a formula for the number of small triangles in a hexagonal array of any given side-length. Their interaction was guided by a sequence of proposals and responses carried out textually in the chat medium. However, the sense of the terms and relationships narrated in the chat were largely instantiated, shared, and investigated through observation of visible features of graphical inscriptions in the whiteboard medium. The mathematical object that was visually co-constructed in the whiteboard was named and described in words within the chat. Finally, a symbolic expression was developed by the group, grounded in the graphic that evolved in the whiteboard and discussed in the terminology that emerged in the chat. The symbolic mathematical result was then posted to the wiki, a third medium within the VMT environment. The wiki is intended for sharing group findings with other groups as part of a permanent archive of work by virtual math teams.

Our case study in this paper demonstrates that it is possible to analyze how math problem solving—and presumably other cognitive achievements—can be carried out by small groups of students. The students can define and refine their own problems to pursue; they can invent their own methods of working; they can use unrestricted vocabulary; they can coordinate work in multiple media, taking advantage of different affordances. Careful attention to the sequentiality of references and responses is necessary to reveal *how* the group coordinated its work and how that work was driven by the reactions of the group members' actions to each other. Only by focusing on the sequentiality of the actions can one see how the visual, narrative, and symbolic build on each other as well as how the actions of the individual students respond to each other. Through these actions, the students co-construct math objects, personal understanding, group agreement, and mathematical results that cannot be attributed to any one individual, but that emerge from the interaction as complexly sequenced.

This analysis illustrates a promising approach for CSCL research to investigate aspects of group cognition that are beyond the reach of alternative methods that systematically ignore the full sequentiality of their data.

The group as the unit of analysis

For methodological reasons, quantitative approaches—such as those reviewed in the next section—generally (a) constrain (scaffold) subject behaviors, (b) filter (code) the data in terms of operationalized variables, and (c) aggregate (count) the coded data. These acts of standardization and reduction of the data eliminate the possibility of observing the details and enacted processes of unique, situated, indexical, sequential, group interaction (Stahl 2006, chap. 10). An alternative form of interaction analysis is needed to explore the organization of interaction that can take place in CSCL settings.

In this paper, we focused on small-group interactions mediated by a multimodal interaction space. Our study differs from similar work in CSCL by our focus on groups larger than dyads whose members are situated outside a controlled lab environment, and by our use of open-ended math tasks where students are encouraged to come up with their own problems. Moreover, we do not impose any deliberate restrictions on the ways students access the features of our online environment or on what they can say. Our main goal is to investigate how small groups of students construe and make use of the "available features" of the VMT online environment to discuss mathematics with peers from different schools outside their classroom setting. In other words, we are interested in studying interactional achievements of small groups in complex computer mediations "in the wild" (Hutchins 1996).

Our interest in studying the use of an online environment with multiple interaction spaces in a more naturalistic use scenario raises serious methodological challenges. In an early VMT study where we conducted a content analysis of collaborative problem-solving activities mediated by a standard text-chat tool in a similar scenario of use, we observed that groups larger than dyads exhibit complex interactional patterns that are difficult to categorize based on a theory-informed coding scheme with a fixed/predetermined unit of analysis (Stahl 2009b, chap. 20). In particular, we observed numerous cases where participants post their messages in multiple chat turns, deal with contributions seemingly out of sequence, and sustain conversations across multiple threads that made it problematic to segment the data into fixed analytic units for categorization. Moreover, coming to agreement on a code assignment for a unit that is defined a priori (e.g., a chat line) turned out to be heavily dependent upon how the unit can be read in relation to resources available to participants (e.g., the problem description) and to prior units (Stahl 2009b, chap. 22). In other words, the sense of a unit not only depends on the semantic import of its constituent elements, but also on the occasion in which it is situated (Heritage 1984). This often makes it possible to apply multiple categories to a given unit and threatens the comparability of cases that are labeled with the same category. More importantly, once the data is reduced to codes and the assignments are aggregated, the complex sequential relationships among the units are largely lost. Hence, the coding approach's attempt to enforce a category to each fixed unit without any consideration to how users sequentially organize their actions in the environment proved to be too restrictive to adequately capture the interactional complexity of chat (Stahl 2009b, chap. 23). Moreover, the inclusion of a shared drawing area in our online environment made the use of a standard coding schema even harder due to increased possibilities for interaction. The open-ended nature of the tasks we use in our study makes it especially challenging to model certain types of actions and to compare them against ideal solutions.

The issue of unit of analysis has theoretical implications. In text chat, it is tempting to take a single posting as the unit to be analyzed and coded, because a participant defined this as a unit by posting it as a message and because the chat software displays it as a visual unit. However, this tends to lead the analyst to treat the posting as a message from the posting individual—that is, as an expression of a thought in the poster's mind, which must then be interpreted in the minds of the post readers. Conversation analysis has argued for the importance of *interactions* among participants as forming more meaningful units for analysis. These consist of sequences of multiple utterances by different speakers; the individual utterances take each other into account. For instance, in a question/answer "adjacency pair," the question elicits an answer and the answer responds to the question. To take a pair of postings such as a question/answer pair as the analytic unit is to treat the interaction within the group as primary. It focuses the analysis at the level of the group rather than the individual. As mentioned, in online text chat, responses are often separated from their referents, so the analysis is more complicated. In general, we find that the important thing is to trace as many references as possible between chat postings or whiteboard actions in order to analyze the interaction of the group as it unfolds (Stahl 2009b, chap. 26). As seen in our case study, it is through the co-construction of a rich nexus of such references that the group weaves its joint problem space.

Analysis at the group unit focuses on the co-construction, maintenance, and progressive refinement of the joint problem space. This is a distinctive analytic task that takes as its data only what is shared by the group. Whatever may go on in the physical, mental, or cultural backgrounds of the individual participants is irrelevant unless it is brought into the group discourse. Because the students know nothing about the gender, age, ethnicity, accent, appearance, location, personality, opinions, grades, or skills of the other participants other than what is mentioned or displayed in the chat interaction, these "factors" from the individual and societal levels can be bracketed out of the group analysis. Survey and interview data is unnecessary; individual learning trajectories are not plotted. The VMT Project has been designed to make available to the analyst precisely what was shared by the student group, and nothing else.

Relatedly, the notion of common ground (see section on grounding below) as an abstract placeholder for registered cumulative facts or pre-established meanings has been critiqued in the CSCL literature for treating meaning as a fixed/denotative entity transcendental to the meaning-making activities of inquirers (Koschmann 2002). The common ground that supports mutual understanding in group cognition or group problem solving is a matter of semantic references that unfold sequentially in the momentary situation of dialog, not a matter of comparing mental contents (Stahl 2006, pp. 353–356). Committing to a reference-repair model (Clark and Marshall 1981) for meaning making falls short of taking into account the dynamic, constitutive nature of meaning-making interactions that foster the process of inquiry (Koschmann et al. 2001).

As we saw in the preceding case study, the understanding of the mathematical structure of the hexagon area did not occur as a mental model of one of the students that was subsequently externalized in the chat and whiteboard and communicated to the other students. It emerged in the discourse media in a way that we could witness as analysts. It consisted of the layering of inscriptions (textual and graphical) that referenced one another. The referential network of group meaning can be observed in the way that deictic and indexical expressions are resolved. The three students each contribute to the progressive development of the shared meaning by responding appropriately to the ongoing state of the discourse. This is a matter of linguistic skill—including ability in discussing mathematical matters—not of articulating mental representations. It is surprising from a rationalist perspective how poor students are at explaining (Stahl 2009b, chap. 26), reproducing (Koschmann and LeBaron 2003), or even recalling (Stahl 2009b, chap. 6) what they did in the group when they are no longer situated in the moment.

Given these analytical and theoretical issues, we opted for an alternative to the approaches reviewed below that involve modeling of actions and correct solution paths or treating shared understanding as alignment of preexisting individual representations and opinions. In this paper, we built on our previous work on referencing math objects in a system with chat and a whiteboard (Stahl 2009b, chap. 17); we presented a "micro-ethnographic" (Streeck and Mehus 2003) case study using interaction analysis (Jordan and Henderson 1995). We focused on the *sequence of actions* in which the group co-constructs and makes use of *semiotic resources* (Goodwin 2000) distributed across dual-interaction spaces to *do* collaborative problem-solving work. In particular, we focused on the joint organization of activities that produce graphical drawings on the shared whiteboard and the ways those drawings are used as resources by actors as they collaboratively work on an open-ended math task. Through detailed analysis at the group unit of analysis, we investigated how actions performed in one workspace inform the actions performed in the other and how the group coordinates its actions across both interaction spaces.

Other approaches in CSCL to analyzing multimodal interaction

Multimodal interaction spaces—which typically bring together two or more synchronous online communication technologies such as text chat and a shared graphical workspace—have been widely used to support collaborative learning activities of small groups (Dillenbourg and Traum 2006; Jermann 2002; Mühlpfordt and Wessner 2005; Soller and Lesgold 2003; Suthers et al. 2001). The way such systems are designed as a juxtaposition of several technologically independent online communication tools carries important interactional consequences for the users. Engaging in forms of joint activity in such online environments requires group members to use the technological features available to them in methodical ways to make their actions across multiple spaces intelligible to each other and to sustain their joint problem-solving work.

In this section we summarize our review (Çakir 2009) of previous studies in the CSCL research literature that focus on the interactions mediated by systems with multimodal interaction spaces to support collaborative work online. Our review is not meant to be exhaustive, but representative of the more advanced analytical approaches employed. We have selected sophisticated analyses, which go well beyond the standard coding-and-counting genre of CSCL quantitative reports, in which utterances are sorted according to a fixed coding scheme and then statistics are derived from the count of utterances in each category. Unlike the simple coding-and-counting studies, the approaches we review attempt to analyze some of the structure of the semantic and temporal relationships among chat utterances and workspace inscriptions in an effort to get at the fabric of common ground in dual-interaction online environments.

The communicative processes mediated by multimodal interaction spaces have attracted increasing analytical interest in the CSCL community. A workshop held at CSCL 2005 specifically highlighted the need for more systematic ways to investigate the unique affordances of such online environments (Dillenbourg 2005). Previous CSCL studies that focus on the interactions mediated by systems with two or more interaction spaces can be broadly categorized under: (1) prescriptive approaches based on models of interaction and (2) descriptive approaches based on content analysis of user actions.

(1) The *modeling approach* builds on the content-coding approach by devising models of categorized user actions performed across multimodal interaction spaces, for example:

   (a) Soller and Lesgold's (2003) use of hidden Markov models (HMM) and
   (b) Avouris et al.'s (2003) object-oriented collaboration analysis framework (OCAF).

In these studies, the online environment is tailored to a specific problem-solving situation so that researchers can partially automate the coding process by narrowing the possibilities for user actions to a well-defined set of categories. The specificity of the problem-solving situation also allows researchers to produce models of idealized solution cases. Such ideal cases are then used as a baseline to make automated assessments of group work and learning outcomes.

(2) The *descriptive approach* informed by content analysis also involves categorization of user actions mediated by multimodal interaction spaces, applying a theoretically informed coding scheme. Categorized interaction logs are then subjected to statistical analysis to investigate various aspects of collaborative work such as:

   (c) The correlation between planning moves performed in chat and the success of subsequent manipulations performed in a shared workspace (Jermann 2002; Jermann and Dillenbourg 2005),
   (d) The relationship between grounding and problem-solving processes across multiple interaction spaces (Dillenbourg and Traum 2006),
   (e) A similar approach based on cultural-historical activity theory (Baker et al. 1999), and
   (f) The referential uses of graphical representations in a shared workspace in the absence of explicit gestural deixis (Suthers et al. 2003).

These studies all focus on the group processes of collaboration, rather than treating it as a mere experimental condition for comparing the individuals in the groups. Also, they employ a content-coding approach to categorize actions occurring in multiple interaction spaces. In most cases, representational features like sentence openers or nodes corresponding to specific ontological entities are implemented in the interface to guide/constrain the possibilities for interaction. Such features are also used to aid the categorization of user actions. The categorization schemes are applied to recorded logs and subjected to statistical analysis to elicit interaction patterns.

The analytic thrust of these studies is to arrive at quantitative results through statistical comparisons of aggregated data. To accomplish this, they generally have to restrict student actions in order to control variables in their studies and to facilitate the coding of student utterances within a fixed ontology. We fear that this unduly restricts the interaction, which must be flexible enough to allow students to invent unanticipated behaviors. The restrictions of laboratory settings make problematic experimental validity and generalization of results to real-world contexts. Even more seriously, the aggregation of data—

grouping utterances by types or codes rather than maintaining their sequentiality—ignores the complexity of the relations among the utterances and actions. According to our analysis, the temporal and semiotic relations are essential to understanding, sharing, and coordinating meaning, problem solving, and cognition. While quantitative approaches can be effective in testing model-based hypotheses, they seem less appropriate both for exploring the problem of interactional organization and for investigating interactional methods, which we feel are central to CSCL theory.

Despite the accomplishments of these studies, we find that their approaches introduce systematic limitations. Interactional analysis is impossible because coherent excerpts from recorded interactions are excluded from the analysis itself. (Excerpts are only used anecdotally, outside of the analysis, to introduce the features of the system to the reader, to illustrate the categorization schemes employed, or to motivate speculative discussion). Moreover, most studies like these involve dyads working on specific problem-solving contexts through highly structured interfaces in controlled lab studies in an effort to manage the complexity of collaboration. The meanings attributed by the researchers to such features of the interface need to be discovered/unpacked by the participants as they put them into use in interaction—and this critical process is necessarily ignored by the methodology. Finally, most of these papers are informed by the psycholinguistic theory of common ground, and are unable to critique it systematically. By contrast—as we shall see in the following section—our analysis of the joint organization of interaction in the case study positions us to understand how the group grounds its shared understanding in interactional terms at the group level.

Grounding through interactional organization

The coordination of visual and linguistic methods (across the whiteboard and chat workspaces) plays an important role in the establishment of common ground through the co-construction of references between items in the different media within the VMT environment. Particularly in mathematics—with its geometric/algebraic dual nature—symbolic terms are often grounded in visual presence and associated visual practices, such as counting or collecting multiple units into a single referent (Goodwin 1994; Healy and Hoyles 1999; Livingston 2006; Sfard 2008; Wittgenstein 1944/1956). The visually present can be replaced by linguistic references to objects that are no longer in the visual field, but that can be understood based on prior experience supported by some mediating object such as a name—see the discussion of mediated memory and of the power of names in thought by Vygotsky (1930/1978, 1934/1986). A more extended analysis of the co-construction of mathematical artifacts by virtual math teams, the complementarity of their visual, semantic, and symbolic aspects, their reliance on pre-mathematical practices and processes of reification into concepts are beyond the scope of this paper and require comparison of multiple case studies (see Çakir 2009). However, for this paper it is important to understand something of how the interactional organization that we have observed here functions to ground the group's understanding of their math object (the hexagonal array) as a shared group achievement.

As implied in the OCAF study (Avouris et al. 2003) mentioned in the previous section, investigating grounding and problem-solving processes in online dual-interaction environments like VMT requires close attention to the relationships among actions performed in multiple interaction spaces. Our case study illustrates some of the practical challenges involved with producing mathematical models that aim to exhaustively capture such relationships. For instance, the hexagonal array that was co-constructed by the team draws upon a triangular grid that is formed by three sets of parallel lines that intersect with each other in a particular way. In other words, these objects are layered on top of each other by the participants to produce a shape recognizable as a hexagon. Despite this combinatoric challenge, a modeling approach can still attempt to capture all possible geometric relationships among these graphical objects in a bottom-up fashion. However, when all chat messages referring to the whiteboard objects are added to the mix, the resulting model may obscure rather than reveal the details of the interactional organization through which group members discuss more complicated mathematical objects by treating a collection of atomic actions as a single entity. Terminology co-constructed in the chat-and-whiteboard environment—like "hexagonal array"—can refer to complexly defined math objects. What is interesting about the student knowledge building is how they aggregate elements and reify them into higher order, more powerful units (Sfard 2008). A model should mirror this rather than to simply represent the elements as isolated.

The challenges involved with the modeling approach are not limited to finding efficient ways to capture all relationships among actions and identifying meaningful clusters of objects. The figurative uses of the graphical objects present the most daunting challenge for such an undertaking. For instance, the team members in our case study used the term "hexagonal array" to refer to a mathematical object implicated in the witnessed production of prior drawing actions. As we have seen in the way the team used this term during their session, "hexagonal array" does not simply refer to a readily available whiteboard illustration. Instead it is used as a *gloss* (Garfinkel and Sacks 1970) to talk about an imagined pattern that grows infinitely and takes the shape illustrated on the whiteboard only at a particular stage. In the absence of a fixed set of ontological elements and constraints on types of actions a user can perform, modeling approaches that aim to capture emergent relationships among semiotic objects distributed across multiple interaction spaces need to adequately deal with the retrospective and prospective uses of language in interaction. Rather than relying upon a generic approach to modeling imposed by the researchers, our ethnographic approach aims to discover the unique "model"—or, better, the specific meaning—that was constructed *by the group* in its particular situation.

In another study discussed earlier, Dillenbourg and Traum (2006) offer the napkin and mockup models in their effort to characterize the relationship between whiteboard and chat spaces. In short, these models seem to describe two use scenarios where one interaction space is subordinated to the other during an entire problem-solving session. The complex relationships between the actions performed across both interaction spaces in our case made it difficult for us to describe the interactions we have observed by committing to only one of these models, as Dillenbourg & Traum did in their study. Instead, we have observed that in the context of an open-ended math task, groups may invoke either type of organization, depending upon the contingencies of their ongoing problem-solving work. For instance, during long episodes of drawing actions where a model of some aspect of the shared task is being co-constructed on the whiteboard (as in our first excerpt), the chat area often serves as an auxiliary medium to coordinate the drawing actions, which seems to conform to the mockup model. In contrast, when a strategy to address the shared task is being discussed in chat (as in the excerpt where the group considered splitting the hexagon into six regions), the whiteboard may be mainly used to quickly illustrate the textual descriptions with annotations or rough sketches, in accordance with the napkin model. Depending on the circumstances of ongoing interaction, participants may switch from one type of organization to another from moment to moment. Therefore, instead of ascribing mockup and napkin models to entire problem-solving sessions, we argue that it would be more fruitful to use these terms as glosses or descriptive categories for types of interactional organizations that group members may invoke during specific episodes of their interaction.

Another provocative observation made by Dillenbourg & Traum is that the whiteboard serves as a kind of shared external memory where group members keep a record of agreed-upon facts. In their study, the dyads were reported to post text notes on the whiteboard to keep track of the information they had discovered about a murder-mystery task. This seems to have led the authors to characterize the whiteboard as a placeholder and/or a shared working memory for the group, where agreed-upon facts or "contributions" in Clark's sense are persistently stored and spatially organized. As Dillenbourg & Traum observed, the scale of what is shared in the course of collaborative problem solving becomes an important issue when a theory operating at the utterance level like contribution theory (Clark and Marshall 1981) is used as an analytic resource to study grounding processes that span a longer period of time. Dillenbourg & Traum seem to have used the notion of persistence to extend common ground across time to address this limitation. In particular, they argued that the whiteboard grounds the solution to the problem itself rather than the contributions made by each utterance. In other words, the whiteboard is metaphorically treated as a physical manifestation of the common ground. We certainly agree with this broadening of the conceptualization of common ground, although we do not see the whiteboard as just a metaphor or externalization of a mental phenomenon. Rather, *common ground is established in the discourse spaces* of text chat and graphical whiteboard. Their differential forms of persistence provide a continuing resource for sharing, modifying, and remembering the group meaning of joint artifacts and products of group cognition.

In our case study, we have observed that the whiteboard does not simply serve as a kind of shared external memory where the group keeps a record of agreed-upon facts, opinions, hypotheses, or conclusions. The shared visible communication media are places where the group does its work, where it cognizes. Ideas, concepts, meanings, and so forth can *subsequently* be taken up by individuals into their personal memories as resources for future social or mental interactions. There is no need to reduce group meaning to identical individual mental contents or to hypothesize a mysterious "group mind" as the location of common ground—the location is the discourse medium, with all its particular affordances and modes of access.

In our sessions, the whiteboard was primarily used to draw and annotate graphical illustrations of geometric shapes, although users occasionally posted textboxes on the whiteboard to note formulas they had found (see Fig. 2 above). While the whiteboard mainly supported visual reasoning—and textual discussion or symbolic manipulation occurred chiefly in the chat stream—actions were carefully, systematically coordinated across the media and integrated within an interactionally organized group-cognitive process. As we have illustrated in our analysis, the fact that there were inscriptions posted on the whiteboard did not necessarily mean that all members immediately shared the same sense of those graphical objects. The group members did considerable interactional work to achieve a shared sense of those objects that was adequate for the purposes at hand. For instance, the crosshatched lines that Qwertyuiop originally drew became increasingly meaningful for the group as it was visually outlined and segmented and as it was discussed in the chat and expressed symbolically.

Hence, the whiteboard objects have a different epistemic status in our case study than in Dillenbourg & Traum's experiment. Moreover, the participants did not deem all the contents of the whiteboard relevant to the ongoing discussion. For instance, Fig. 2 above shows a snapshot of the entire whiteboard as the team was discussing the hexagonal pattern problem. The figure shows that there are additional objects in the shared scene like a blue hypercube and a 3-D staircase, which are remnants of the group's prior problem-solving work. Finally, the sense of previously posted whiteboard objects may be modified or become evident as a result of current actions (Suchman 1990).

In other words, group members can not only reuse or reproduce drawings, but they can also make subsequent sense of those drawings or discard the ones that are not deemed relevant anymore. Therefore, the technologically extended notion of common ground as a placeholder for a worked-out solution suffers from the same issues stated in Koschmann and LeBaron's (2003) critique of Clark's theory. As an abstract construct transcendental to the meaning-making practices of participants, the notion of common ground obscures rather than explains the ways the whiteboard is used as a resource for collaborative problem solving.

Instead of using an extended version of common ground as an analytical resource, we frame our analysis using the notion of "indexical ground of deictic reference," which is a notion we appropriated from linguistic anthropology (Hanks 1992). In face-to-face interaction, human action is built through the sequential organization of not only talk but also coordinated use of the features of the local scene that are made relevant via bodily orientations, gesture, eye gaze, and so forth. In other words, "human action is built through simultaneous deployment of a range of quite different kinds of semiotic resources" (Goodwin 2000, p. 1489). Indexical terms and referential deixis play a fundamental role in the way these semiotic resources are interwoven in interaction into a coherent whole.

Indexical terms are generally defined as expressions whose interpretation requires identification of some element of the context in which it was uttered, such as who made the utterance, to whom it was addressed, when and where the utterance was made (Levinson 1983). Because the sense of indexical terms depends on the context in which they are uttered, indexicality is necessarily a relational phenomenon. Indexical references facilitate the mutually constitutive relationship between language and context (Hanks 1996). The basic communicative function of indexical-referentials is "to individuate or single out objects of reference or address in terms of their relation to the current interactive context in which the utterance occurs" (Hanks 1992, p. 47).

The specific sense of referential terms such as *this*, *that*, *now*, *here* is defined locally by interlocutors against a shared indexical ground. Conversely, the linguistic labels assigned to highlighted features of the local scene shapes the indexical ground. Hence, the indexical ground is not an abstract placeholder for a fixed set of registered contributions. Rather, it signifies an emergently coherent field of action that encodes an interactionally achieved set of background understandings, orientations, and perspectives that make references intelligible to interlocutors (Zemel et al. 2008).

Despite the limitations of online environments for supporting multimodality of embodied interaction, participants make substantial use of their everyday interactional competencies as they appropriate the features of such environments to engage with other users. For instance, Suthers et al.'s (2003) study reports that deictic uses of representational proxies play an important role in the interactional organization of online problem-solving sessions mediated by the Belvedere system. The authors report that participants in the online case devised mechanisms that compensate for the lack of gestural deixis with alternative means, such as using verbal deixis to refer to the most recently added text nodes and visual manipulation of nodes to direct their partner's attention to a particular node in the shared argument map.

In contrast to the Belvedere system, VMT offers participants additional resources such as an explicit referencing mechanism, a more generic workspace that allows producing and annotating drawings, and an awareness feature that produces a sense of sequentiality by embedding indicators for drawing actions in the sequence of chat postings. Our case study shows that despite the online situation's lack of the familiar resources of embodied interaction, team members can still achieve a sense of shared access to the meaningful objects displayed in the dual-interaction spaces of the VMT environment. Our analysis indicates that coherence among multiple modalities of an online environment like VMT is achieved through group members' development and application of shared methods for using the features of the system to coordinate their actions in the interface.

Through coordinated use of indexical-referential terms and highlighting actions, team members help each other to literally "see" the objects implicated in the shared visual field (Goodwin 1994) and to encode them with locally specified terminology for subsequent use. They demonstrate how to "read" graphical as well as textual objects through the way the objects are built up sequentially and are spatially arranged in relation to each other through sequences of actions. The deictic references that link chat messages to features of graphical inscriptions and to prior chat content are instrumental in the sequential achievement of indexical symmetry, intersubjectivity, or common ground.

Sequential analysis of the joint organization of interaction

To sum up, the focus of our ethnomethodological inquiry is directed toward documenting how a virtual team achieved intersubjectivity and coherence among their actions in an online CSCL environment with multiple interaction spaces. We looked at the moment-to-moment details of the practices through which participants organize their chat utterances and whiteboard actions as a coherent whole in interaction—a process that is central to CSCL. We observed that referential practices enacted by the users are essential, particularly in the coordinated use of multimodalities afforded by environments like VMT. The referential uses of available features are instrumental not only in allocating other members' attention to specific parts of the interface where relevant actions are being performed, but also in the achievement of reciprocity (intersubjectivity, common ground, shared understanding, group cognition) among actions in the multiple interaction spaces, and hence, a sense of sequential organization across the spaces.

In our case study, we have seen the establishment of an indexical ground of deictic references co-constructed by the group members as an underlying support for the creation and maintenance of their joint problem space. We have seen that nexus of references created interactionally as group members propose, question, repair, respond, illustrate, make visible, supply symbols, name, and so forth. In the VMT dual-media environment, the differential persistence, visibility, and mutability of the media are consequential for the interaction. Group members develop methods of coordinating chat and drawing activities to combine visual and conceptual reasoning by the group and to co-construct and maintain an evolving shared indexical ground of their discourse.

In this paper, we have reconceptualized the problem of common ground from an issue of sharing mental representations to a practical matter of being able to jointly relate semiotic objects to their indexed referents. The references do not reside in the minds of particular actors, but have been crafted into the presentation of the chat postings and drawing inscriptions through the details of wording and sequential presentation. The references are present in the data as affordances for *understanding* by group participants as well as by analysts (Stahl 2006, chap. 17). The *meaning* is there in the visual presentation of the communication objects and in the network of interrelated references (Stahl 2007), rather than in mental re-presentations of them. The understanding of the references is a matter of normally tacit social practice, rather than of rationalist explicit deduction. The references can be explicated by analysis, but only if the structure of sequentiality and indexicality is preserved in the data analysis and only if the skill of situated human understanding is applied.

In our case study of an 18-min excerpt taken from a four-hour group chat, three students construct a diagram of lines, triangles, and hexagons, propose a math pattern problem, analyze the structure of their diagram, and derive an algebraic formula to solve their problem. They propose their own creative problem about mathematical properties; gradually construct a complex mathematical object; explore related patterns with visual, narrative, and symbolic means; express wonder; gain mathematical insight; and appreciate their achievement. They do this by coordinating their whiteboard and chat activities in a synchronous online environment. Their accomplishment is precisely the kind of educational math experience recommended by mathematicians (Livingston 2006; Lockhart 2008; Moss and Beatty 2006). It was not a mental achievement of an individual, but a group accomplishment carried out in computer-supported discourse. By analyzing the sequentiality and indexicality of their interactions, we explicated several mechanisms of the group cognition by which the students coordinated the group meaning of their discourse and maintained an effective joint problem space.

The coordination of visual and textual realizations of the mathematical objects that the students co-construct provides a grounding of the algebraic formulas the students jointly derive using the line drawings that they inspect visually together. As the students individualize this experience of group cognition, they can develop the deep understanding of mathematical phenomena that comes from seeing the connections among multiple realizations (Sfard 2008; Stahl 2008). Our case study does not by any means predict that all students can accomplish similar results under specific conditions, but merely demonstrates that such group cognition is possible within a synchronous CSCL setting and that a fine-grained sequential analysis of interaction can study how it is collaboratively accomplished.

Acknowledgment The reviews coordinated by Dan Suthers helped us to structure this paper more clearly. Some of the larger methodological, technological, and pedagogical issues the reviewers raised are addressed in (Stahl 2009b), which lists the VMT research team members. This paper is a result of the team's group cognition. Access to the complete data using the VMT Replayer is available by emailing the authors.

References

Avouris, N., Dimitracopoulou, A., & Komis, V. (2003). On analysis of collaborative problem solving: an object-oriented approach. *Computers in Human Behavior, 19*, 147–167.

Baker, M., Hansen, T., Joiner, R., & Traum, D. (1999). The role of grounding in collaborative learning tasks. In P. Dillenbourg (Ed.), *Collaborative learning: Cognitive and computational approaches* (pp. 31–63). Oxford, UK: Pergamon.

Çakir, M. P. (2009). *How online small groups co-construct mathematical artifacts to do collaborative problem solving*. Philadelphia, PA, USA: Unpublished Dissertation, Ph.D., College of Information Science and Technology, Drexel University.

Clark, H., & Brennan, S. (1991). Grounding in communication. In L. Resnick, J. Levine & S. Teasley (Eds.), *Perspectives on socially-shared cognition* (pp. 127–149). Washington, DC: APA.

Clark, H. H., & Marshall, C. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. Weber & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). New York, NY: Cambridge University Press.

Coulon, A. (1995). *Ethnomethodology*. Thousand Oaks, CA: Sage.

Dillenbourg, P. (2005). Dual-interaction spaces. In T. Koschmann, D. D. Suthers & T.-W. Chan (Eds.), *Computer-supported collaborative learning 2005: The next ten years! (Proceedings of CSCL 2005)*. Taipei, Taiwan: Mahwah, NJ: Erlbaum.

Dillenbourg, P., Baker, M., Blaye, A., & O'Malley, C. (1996). The evolution of research on collaborative learning. In P. Reimann & H. Spada (Eds.), *Learning in humans and machines: Towards an interdisciplinary learning science* (pp. 189–211). Oxford, UK: Elsevier.
Dillenbourg, P., & Traum, D. (2006). Sharing solutions: persistence and grounding in multimodal collaborative problem solving. *Journal of the Learning Sciences, 15*(1), 121–151.
Garcia, A., & Jacobs, J. B. (1998). The interactional organization of computer mediated communication in the college classroom. *Qualitative Sociology, 21*(3), 299–317.
Garcia, A., & Jacobs, J. B. (1999). The eyes of the beholder: understanding the turn-taking system in quasi-synchronous computer-mediated communication. *Research on Language and Social Interaction, 34*(4), 337–367.
Garfinkel, H., & Sacks, H. (1970). On formal structures of practical actions. In J. Mckinney & E. Tiryakian (Eds.), *Theoretical sociology: Perspectives and developments* (pp. 337–366). New York, NY: Appleton-Century-Crofts.
Goodwin, C. (1994). Professional vision. *American Anthropologist, 96*(3), 606–633.
Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics, 32*, 1489–1522.
Goodwin, C., & Heritage, J. (1990). Conversation analysis. *Annual Review of Anthropology, 19*, 283–307.
Hanks, W. (1992). The indexical ground of deictic reference. In C. Goodwin & A. Duranti (Eds.), *Rethinking context: Language as an interactive phenomenon*. Cambridge, UK: Cambridge University Press.
Hanks, W. (1996). *Language and communicative practices*. Boulder, CO: Westview.
Healy, L., & Hoyles, C. (1999). Visual and symbolic reasoning in mathematics: making connections with computers. *Mathematical Thinking and Learning, 1*(1), 59–84.
Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge, UK: Polity.
Hutchins, E. (1996). *Cognition in the wild*. Cambridge, MA: MIT.
Jermann, P. (2002). *Task and interaction regulation in controlling a traffic simulation.* Paper presented at the Computer support for collaborative learning: Foundations for a CSCL community. Proceedings of CSCL 2002, Boulder, CO. Proceedings pp. 601–602.
Jermann, P., & Dillenbourg, P. (2005). Planning congruence in dual spaces. In T. Koschmann, D. D. Suthers & T.-W. Chan (Eds.), *Computer-supported collaborative learning 2005: The next ten years! (Proceedings of CSCL 2005)*. Taipei, Taiwan: Mahwah, NJ: Erlbaum.
Jordan, B., & Henderson, A. (1995). Interaction analysis: foundations and practice. *Journal of the Learning Sciences, 4*(1), 39–103. Retrieved from http://lrs.ed.uiuc.edu/students/c-merkel/document4.HTM.
Koschmann, T. (2002). Dewey's contribution to the foundations of CSCL research. In G. Stahl (Ed.), *Computer support for collaborative learning: Foundations for a CSCL community: Proceedings of CSCL 2002*, pp. 17–22. Boulder, CO: Erlbaum.
Koschmann, T., & LeBaron, C. (2003). *Reconsidering common ground: Examining clark's contribution theory in the operating room.* Paper presented at the European Computer-Supported Cooperative Work (ECSCW '03), Helsinki, Finland. Proceedings pp. 81–98.
Koschmann, T., LeBaron, C., Goodwin, C., & Feltovich, P. J. (2001). Dissecting common ground: Examining an instance of reference repair. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the twenty-third annual conference of the cognitive science society* (pp. 516–521). Mahwah, NJ: Erlbaum.
Koschmann, T., Stahl, G., & Zemel, A. (2007). The video analyst's manifesto (or the implications of Garfinkel's policies for the development of a program of video analytic research within the learning sciences). In R. Goldman, R. Pea, B. Barron & S. Derry (Eds.), *Video research in the learning sciences* (pp. 133–144). Erlbaum: Mahway, NJ. Retrieved from http://GerryStahl.net/publications/journals/manifesto.pdf.
Levinson, S. (1983). *Pragmatics*. Cambridge, UK: Cambridge University Press.
Livingston, E. (2006). Ethnomethodological studies of mediated interaction and mundane expertise. *The Sociological Review, 54*(3).
Lockhart, P. (2008). Lockhart's lament. *MAA Online, 2008*(March). Retrieved from http://www.maa.org/devlin/devlin_03_08.html.
Moss, J., & Beatty, R. A. (2006). Knowledge building in mathematics: supporting collaborative learning in pattern problems. *International Journal of Computer-Supported Collaborative Learning (ijCSCL), 1*(4), 441–466.
Mühlpfordt, M., & Stahl, G. (2007). The integration of synchronous communication across dual interaction spaces. In C. Chinn, G. Erkens, & S. Puntambekar (Eds.), *The proceedings of CSCL 2007: Of mice, minds, and society (CSCL 2007)*. New Brunswick, NJ. Retrieved from http://GerryStahl.net/vmtwiki/martin.pdf.
Mühlpfordt, M., & Wessner, M. (2005). Explicit referencing in chat supports collaborative learning. In T. Koschmann, D. D. Suthers & T.-W. Chan (Eds.), *Computer-supported collaborative learning 2005: The next ten years! (Proceedings of CSCL 2005)* (pp. 460–469). Taipei, Taiwan: Mahwah, NJ: Erlbaum.
O'Neill, J., & Martin, D. (2003). *Text chat in action.* Paper presented at the ACM Conference on Groupware (GROUP 2003), Sanibel Island, FL.

Psathas, G. (1995). *Conversation analysis: The study of talk-in-interaction*. Thousand Oaks, CA: Sage.
Roschelle, J., & Teasley, S. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69–197). Berlin, Germany: Springer Verlag.
Sacks, H. (1962/1995). *Lectures on conversation*. Oxford, UK: Blackwell.
Sarmiento, J., & Stahl, G. (2008). *Extending the joint problem space: Time and sequence as essential features of knowledge building.* Paper presented at the International Conference of the Learning Sciences (ICLS 2008), Utrecht, Netherlands. Retrieved from http://GerryStahl.net/pub/icls2008johann.pdf.
Sarmiento-Klapper, J. W. (2009). *Bridging mechanisms in team-based online problem solving: continuity in building collaborative knowledge.* Philadelphia, PA, USA: Unpublished Dissertation, Ph.D., College of Information Science and Technology, Drexel University.
Scardamalia, M., & Bereiter, C. (1996). Computer support for knowledge-building communities. In T. Koschmann (Ed.), *CSCL: Theory and practice of an emerging paradigm* (pp. 249–268). Hillsdale, NJ: Erlbaum.
Schegloff, E., & Sacks, H. (1973). Opening up closings. *Semiotica, 8*, 289–327.
Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses and mathematizing*. Cambridge, UK: Cambridge University Press.
Soller, A., & Lesgold, A. (2003). *A computational approach to analyzing online knowledge sharing interaction.* Paper presented at the 11th International Conference on Artificial Intelligence in Education, AI-ED 2003, Sydney, Australia. Proceedings pp. 253–260. Amsterdam: IOS.
Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT. Retrieved from http://GerryStahl.net/mit/.
Stahl, G. (2007). *Meaning making in CSCL: Conditions and preconditions for cognitive processes by groups.* Paper presented at the international conference on Computer-Supported Collaborative Learning (CSCL '07), New Brunswick, NJ: ISLS. Retrieved from http://GerryStahl.net/pub/cscl07.pdf.
Stahl, G. (2008). Book review: exploring thinking as communicating in CSCL. *International Journal of Computer-Supported Collaborative Learning (ijCSCL), 3*(3), 361–368.
Stahl, G. (2009a). *For a science of group interaction.* Paper presented at the GROUP 2009, Sanibel Island, FL.
Stahl, G. (Ed.). (2009b). *Studying virtual math teams*. New York, NY: Springer. Computer-supported collaborative learning book series, vol 11 Retrieved from http://GerryStahl.net/vmt/book.
Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), *Cambridge handbook of the learning sciences* (pp. 409–426). Cambridge, UK: Cambridge University Press. Retrieved from http://GerryStahl.net/cscl/CSCL_English.pdf in English, http://GerryStahl.net/cscl/CSCL_Chinese_simplified.pdf in simplified Chinese, http://GerryStahl.net/cscl/CSCL_Chinese_traditional.pdf in traditional Chinese, http://GerryStahl.net/cscl/CSCL_Spanish.pdf in Spanish, http://GerryStahl.net/cscl/CSCL_Portuguese.pdf in Portuguese, http://GerryStahl.net/cscl/CSCL_German.pdf in German, http://GerryStahl.net/cscl/CSCL_Romanian.pdf in Romanian.
Streeck, J., & Mehus, S. (2003). Microethnography: The study of practices. In K. F. R. Sanders (Ed.), *Handbook of language and social interaction*. Mahway, NJ: Erlbaum.
Suchman, L. A. (1990). Representing practice in cognitive science. In M. Lynch & S. Woolgar (Eds.), *Representation in scientific practice*. Cambridge, MA: MIT.
Suthers, D., Connelly, J., Lesgold, A., Paolucci, M., Toth, E., Toth, J., et al. (2001). Representational and advisory guidance for students learning scientific inquiry. In K. D. Forbus & P. J. Feltovich (Eds.), *Smart machines in education: The coming revolution in educational technology* (pp. 7–35). Menlo Park: AAAI.
Suthers, D., Girardeau, L., & Hundhausen, C. (2003). Deictic roles of external representations in face-to-face and online collaboration. In B. Wasson, S. Ludvigsen & U. Hoppe (Eds.), *Designing for change in networked learning environments, Proceedings of the international conference on computer support for collaborative learning 2003* (pp. 173–182). Dordrecht: Kluwer Academic.
Teasley, S. D., & Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools* (pp. 229–258). Mahwah, NJ: Erlbaum.
ten Have, P. (1999). *Doing conversation analysis: A practical guide*. Thousand Oaks, CA: Sage.
Vygotsky, L. (1930/1978). *Mind in society*. Cambridge, MA: Harvard University Press.
Vygotsky, L. (1934/1986). *Thought and language*. Cambridge, MA: MIT.
Watson, A., & Mason, J. (2005). *Mathematics as a constructive activity: Learners generating examples*. Mahwah, NJ: Erlbaum.
Wittgenstein, L. (1944/1956). *Remarks on the foundations of mathematics*. Cambridge, MA: MIT.
Zemel, A., Koschmann, T., LeBaron, C., & Feltovich, F. (2008). "What are we missing?" Usability's indexical ground. *Computer Supported Cooperative Work, 17*, 63–85.