
Making Group Cognition Visible

How can researchers observe group learning and group cognition; how can these phenomena be made visible for analysis? This chapter addresses the core methodological question for CSCL, borrowing heavily from Garfinkel.

The researcher's interpretive perspective must first be distinguished from, and then be related to, those of the individual group members, the group as a whole and designers of any technical, pedagogical or social innovations. Scientific interpretation of group meaning can then proceed in accordance with ethnomethodology's principles that the data for such analysis is everywhere, visible, grounded, meaningful and situated.

The results reveal the structure of the self-organization of group discourse. The discourse is the embodiment of group cognitive processes, and the analysis of that discourse makes the group's learning and meaning making visible and comprehensible.

How is it possible to rigorously analyze collaborative group meaning making in specific case studies? In this chapter we will address the problem of defining a methodology for making group meaning visible to researchers. We will guide this inquiry with two specific examples. The primary example will be the analysis of mediated collaboration in the `SimRocket` discussion in chapter 12 of part II. In addition, we will use the data from the Virtual Math Teams (VMT) project presented in chapter 17; this will provide an additional example in which the discourse is computer mediated. We will further analyze this data in chapter 21. The `SimRocket` collaboration was face-to-face and videotaped; it was mediated by the computer simulation of model rockets, including the list of rocket components. The VMT data will consist primarily of chat logs taken directly from the computer software that mediates the online collaboration. The question for this chapter is how we can understand a methodology for analysis of these two kinds of cases within the theoretical framework that is being developed in part III.

Perspectives on Collaboration

First, let us be perfectly clear that the kind of analysis we are talking about is necessarily interpretive. The data is language used by people in specific settings—it is not the kind of thing one can simply count up without worrying about what the counted objects meant to the people who uttered and responded to them. Interpretation is perspectival. We argued in chapter 4 that interpretation is necessarily conducted from one interpretive perspective or another. For instance, the perspective from which we are analyzing as researchers is different from the participants' discourse perspectives that we are analyzing. In order to understand the analysis as a process of interpretation, it is important to distinguish the various interpretive perspectives involved:

1. *Individual members* of the group interpret each other's words and behavior as active participants during the live event of collaboration.
2. *Small groups* of collaborating people construct group meanings and knowledge artifacts through the interaction of contributions from their members.
3. *Communities of practice* preserve and disseminate meanings and artifacts.
4. *Collaboration researchers* interpret the behavior of the group and its members by studying data derived from the event, such as video clips and chat logs.
5. *Educational innovators* who are interested in the design of technical or pedagogical interventions draw design consequences from the analyses of the researchers.

Accordingly, we shall distinguish the following five interpretive perspectives in our discussion of analysis methodology: (1) individual group members, (2) the group as a whole, (3) communities of practice providing socio-cultural context, (4) researchers studying the communication and collaboration and (5) designers creating new forms of software, innovative pedagogy or other social practices for future group members.

Let us consider our central example of analysis. In a moment of collaboration lasting several seconds in a middle-school classroom, a small group of students learned something about the conduct of scientific experimentation using the *SimRocket* list artifact. The students made this knowledge visible for their group, repairing confusions and establishing a shared understanding. The micro discourse analysis of this moment in chapters 12 and 13 illustrated the complexity of collaborative learning and of its analysis.

To make learning visible as researchers, we deconstructed the references within the discourse. Thus, we conducted the analysis from the perspective of researchers and our unit of analysis was the group as a whole. The meaning that the group constructed was analyzed as constituting a network of semantic references within the group interaction, rather than as mental representations of individual group members. No assumptions about mental states or representations were required or relevant to the researcher's analysis. Collaborative learning was viewed as the interactive construction of this referential network. The group's shared understanding consisted in the alignment of utterances, evidencing agreement concerning their referents.

The list artifact was a focus of the student discourse. Viewed within its activity system, learning is a social process in which artifacts—whether physical, digital or linguistic—play central roles. Artifacts like the *SimRocket* software must be understood from all five perspectives: their designers, their users as individuals, their group users, the broader community of stakeholders and their researchers.

As meaningful objects in the world, artifacts, by definition, both provide persistence across the communities and require interpretation by each community. The artifacts are *boundary objects* (Bowker & Star, 2000) that span different communities or cross the boundaries between them and thereby permit understanding of one from the position or perspective of the other. The design community designs into the artifact meaningful affordances that must be properly understood in practice by the user communities and their group and individual members. To evaluate the success of this undertaking, the research community must interpret the designed affordances and also interpret the users' practical understandings of these.

In chapter 13 particularly, we tried to understand the indexical references to the list artifact in the group discourse—a set of references that was particularly hard to understand from a superficial reading of the transcript. We found that the students were

engaged in *making visible* to each other the structure of references within their discourse that had become problematic for them as a group engaged in collaborative learning within a classroom activity structure. In making their learning visible to themselves, they made it visible to us as well. Furthermore, they made visible the central affordance of the artifact, which had until then eluded them and caused their group confusion. The group of students, as a whole, systematically constructed a shared understanding by making increasingly explicit the references from their discourse that had created confusion when different students had constructed divergent interpretations.

The world, situation or activity structure in which the group of students operates consists of a shared network of references among words and artifacts. To design new artifacts for these worlds, designers must understand the nature of these referential networks, build artifacts that fit into and extend these networks in pedagogically desirable ways, and provide tasks and social practices that will lead students to incorporate the artifact's new references meaningfully into their shared understandings. Researchers who understand this process can analyze the artifact affordances and the situated student discourse to assess the effectiveness of collaboration technologies.

Computational artifacts such as scientific simulations, productivity software, organizational knowledge repositories and educational systems are designed by one community (e.g., software developers, educators, domain experts or former employees) for use by another (end-users, students, novices or future employees). The two communities typically operate within contrasting cultures; their shared artifacts must cross cultural boundaries to be effective. Diversity among these interacting communities of practice leads to many of the same issues and misunderstandings as cultural diversity among traditional communities.

A computational artifact embodies meaning in its design, its content and its modes of use. This meaning originates in the goals, theories, history, assumptions, tacit understandings, practices and technologies of the artifact's design community. A user community must activate an understanding of the artifact's meaning within their own community practices and cultural-historical contexts.

Clarifying the different perspectives and their associated communities sheds light on the distinction between group meaning and individual interpretation outlined especially in chapter 16. Meaning is associated with the small-group unit of analysis and is shared within the group against the cultural background of the group's larger community. Interpretation is associated with the individual unit of analysis and takes place against the background understanding of the individuals. Both units can be subject to analysis from the researcher's perspective, possibly independently of the knowledge-constitutive human interests (Habermas, 1965/1971) and goals of the designer perspective. It is in this sense—i.e., for the researcher—that meaning is constructed by small groups, within their discourse communities, and is interpreted by individuals from their personal perspectives, situated in their current activity structures.

Given the diversity between the design and user communities, the question arises: how can the meaning embodied in a computational artifact be activated with sufficient continuity that it fulfills its intended function? A further question for us as researchers is how we as members of a third community can assess the extent to which the designers' intentions (for better or worse) were achieved in the students' accomplishments.

Chapter 13 investigated a process of meaning-activation of a computational artifact through an empirical approach: It conducted a micro-ethnographic analysis of an interaction among middle school students learning how to isolate variables in a computer simulation. The analytic affordances (paired configurations) designed into the computational simulation of rocket launches were activated through the involvement of the students in a specific project activity. Their increasing understanding of the artifact's meaning structure was achieved in group discourse situated within their artifact-centered activity.

This micro-ethnographic analysis is a scientific enterprise, like viewing under a microscope the world within a drop of water, a world that is never seen while crossing the ocean by boat. We tried to uncover general *structures* of the interaction that would be applicable to other cases and that thereby contribute to a theoretical understanding of collaboration. The conversational structures of small-group collaboration are different from those of two-person dialog commonly analyzed by conversation analysts, and this has implications for the theory of collaboration.

This approach to studying collaboration differs radically from both traditional educational research and from quantitative studies in CSCL (see chapter 10), both of which can produce useful complementary findings. Experiments in the Thorndikian educational research tradition focus on pre- and post-test behaviors, inferring from changes what kinds of learning took place in between. Such a methodology is the direct consequence of viewing learning as an internal individual mental process that cannot directly be observed (Koschmann, 2002a). However, if we postulate learning to be a social process, then the conditions are very different. In fact, it is not only necessary for the participants in a collaboration to make their evolving understandings *visible* to each other, this is the very essence of collaborative interaction. As we saw in chapter 12, when the evolving learning of the group is not displayed in a coherent manner, everyone's efforts become directed to producing an evident and mutually understood presentation of shared knowledge. That is, in the breakdown case, the structures that are normally invisible suddenly appear as matters of the utmost concern to the participants, who then make explicit and visible to one another the meaning that their utterances have for them. As researchers who share a cultural literacy with the participants, we can take advantage of such displays to formulate and support our analyses.

Making Learning Visible

In the transcript of chapter 12, the teacher provides efficient guidance by directing attention to the list artifact (1:21:53), defining criteria of sameness and difference (1:22:00), and then allowing the students to solve the task collaboratively (1:22:04). Brent points the way with a bold gesture to what already exists in the list artifact (the descriptions of rockets 1 and 2) as the solution. Jamie clarifies how to take this as the solution. Through a sequence of brief, highly interactive turns, the students collaboratively move from treating the list as inadequate, irrelevant and uninteresting to seeing it as holding the key to solving the group task. The sequence ends with a sense of consensus and collaborative accomplishment. In addition to a solution to the nose cone problem, the group has articulated, accepted and put into conversational practice a terminology for discussing sameness, difference, comparison, etc.

By making explicit the references that grant meaning to the discourse (“one and two”), the students made visible to each other the understanding that was being expressed in the interactions. In particular, they made visible the elliptical, indexical and projective references that had become confused. As researchers, we can take advantage of what the participants made visible to each other to also see what was meant and learned as long as we stand within a shared interpretive horizon with them (Gadamer, 1960/1988). Methodologically, our access to these displays is ensured to the extent that we share membership in the culture of understanding that the participants themselves share. For instance, we are native speakers of English, have experienced middle school classroom culture in America, have a lay understanding of rockets, but may not be privy to the latest teen pop culture or the local lore of the particular classroom, so we can legitimately interpret much but perhaps not all of what goes on. Intersubjective validity, the analog of inter-rater reliability, is established by our developing interpretations of the data within group data sessions and presenting those interpretations in seminars and conferences of peers, where our interpretations must be accepted as plausible from the perspectives of a number of researchers.

It is considerably harder to interpret what learning took place in the collaborative moment than in most of the rest of the three-hour session. When the dialog format between a teacher and one student dominates (as it did in much of the remaining time), one can assume—unless there is evidence to the contrary—that learning has taken place for the student (if not necessarily for the whole class) if the student’s response to the teacher’s question has been evaluated as appropriate by the teacher. One basically follows the teacher’s displayed interpretation of what is unfolding, assigning learning to students who he indicates have responded appropriately to his questions. In a collaborative moment, there is no authority guiding, structuring and evaluating the interaction. Deeper interpretation is required to determine what takes place at all, let alone who learns what, when, where and how. In a CSCL setting, where, for instance, many students may be interacting autonomously within a threaded discussion system on the Internet, one must rely on an analysis of student discourse that has a many-to-many structure rather than having all interaction go through the teacher. The potential here is great because learning can overcome the teacher bottleneck and allow much higher levels of student participation in knowledge-building discourse. The problem is how to assess what learning is taking place.

The factors that have in cases of individual learning been taken to be hidden in mental representations can in cases of collaborative learning be taken to be visible in the discourse. The meaning of utterances—even in elliptical, indexical and projective utterances—can be rigorously interpreted on the basis of interaction data such as digital video or computer chat logs. Learning—now viewed at the small-group unit of analysis—can be taken to be a characteristic of the discourse itself. In addition to the group’s shared understanding, however, one can also determine the interpretive perspectives of the individual members, particularly in cases where there are breakdowns of the shared understanding, individual interpretations diverge and the group members must make things explicit. The question now is how to specify a methodology for making the group meaning-making process visible for researchers.

Video Analysis

We propose adopting a methodology to analyze collaborative interaction called *video analysis* (Heath, 1986). It is called this because it has been largely developed through the mediation of digital video. However, it is also applicable to the analysis of collaborative interactions where traces of the discourse are preserved in other forms sufficiently detailed to allow fine-grained micro-analysis, such as comprehensive computer chat logs.

This methodology is based largely on a tradition of interaction analysis (Jordan & Henderson, 1995) that is popular among communication scientists and anthropologists. Its roots are perhaps most extensively elaborated under the rubrics of ethnomethodology (Garfinkel, 1967, 2002; Heritage, 1984) and conversation analysis (Psathas, 1995; Sacks, 1992; ten Have, 1999). Ethnomethodology (EM) is a discipline that focuses on the procedures (i.e., “methods”) that participants (i.e., “members”) use in making sense of their own social actions and the actions of others. Conversation Analysis (CA) is an area of specialization within EM that focuses specifically on the procedures participants employ in competently producing conversation. It provides a rigorous methodology for studying participants’ sense-making practices in the classroom. By studying the sense-making practices of students and teachers, we can document what an instructional innovation means in interactional terms. The pioneers in these fields have focused on discovering the structures of communication (such as turn-taking), rather than applying their methods to practical ends, like evaluating learning and designing curricular innovations. So, we are borrowing their tools and adapting them within a very different scientific endeavor to the extent that we use these analyses to guide the design of new collaboration technologies and practices.

The method we are recommending is an interpretive (hermeneutic) one. This does not make it subjective. On the contrary, we are interested in analyzing the intersubjective meanings that we find in the physical and visible video or chat record, rather than hypothesizing about what may have taken place in subjective individual minds. Perhaps the hardest thing for newcomers to get used to in CA is the method’s strictures against speculating about what participants were “thinking” when they interacted in certain observable ways. The method relies on the fact that the participants, in interacting with each other, were displaying for each other in visible ways (many of which could be captured on video) words and gestures that made sense to both the actor and the other participants. The record of the interaction typically contains numerous clues as to just what sense this was. Subsequent responses of the participants “take up” this meaning in specific ways. Sometimes it turns out that the meaning of some utterance to the speaker and its meaning to the listener were at odds; this difference becomes visible when the conversation turns to visibly repair the misunderstanding. When no evidence of a misunderstanding appears, the analyst can safely assume that for all practical purposes of that interaction the participants had the same understanding of the interaction. The method of analysis is at pains to ensure that the analyst comes to the same understanding as the participants, given relatively similar access to the same utterances as the participants shared.

According to constructivism, learning is a process of constructing new meanings. But, unlike much constructivism, we do not assume that meaning exists only in individual human minds (see chapter 16). The world is full of meaningful things. Most gestures and

utterances that people make are meaningful—and their meanings are necessarily visible to other people—otherwise they would not be effective means of communication.

When one practices interaction analysis for awhile it becomes clear that it is not necessary to interpret meaningful human actions as the result of premeditated, fully worked-out plans in their heads (Suchman, 1987). People just interact and respond to each other on the spot. They may sometimes silently rehearse little speeches in advance of saying them, control what comes out or reflect upon what they said quickly so they can retrospectively give an account. But these mechanisms seem to be secondary phenomena. They are not at all trustworthy accounts of what people meant or why they said something. In a deep sense, “actions speak louder” than retroactive words. It may not be so bad that analysts cannot read people’s minds—their visible actions are more meaningful. If learning takes place in an interaction, we should be able to observe it by analyzing an adequate record of the interaction. It should show in changes in the way that participants use words, in how they build on each others’ utterances, in their expressions, gazes, postures, expressive noises, in how they interact differently in similar circumstances later.

Learning is subtle. It rarely expresses itself in syntactically perfect complete propositions, like one would think based on textbook presentations of knowledge. It is more likely to reveal itself in how the learner gradually starts to use a term with increasing meaning or begins to approach a problem with greater familiarity. Learning is paradoxical; children acquire vocabulary at an incredible rate, but they only have a glimmer of what a new word means. Learning is situated; someone might be able to use a new resource in the context where it was learned, but not yet elsewhere. Analysts can think that they saw visible learning there, but not be sure what its limits are. Discourse is ambiguous; what is said is often open to multiple consistent interpretations. This opens a creative space in which participants can choose among options for proceeding, and it softens interpersonal commitments to avoid potentially embarrassing social consequences. Analysts must rely on how a given utterance was taken up by other participants—and it still may not be possible to pin down a reading of the utterance with much certainty.

If learning is a process of making new meanings, then instruction consists of forms of interactional practice that foster this process. The instruction’s job is to guide the learner or learners in constructing new meaning—that is, in understanding the meaning that is visibly co-constructed in an interaction. The instruction’s job is ultimately to facilitate the learner’s acquisition of the ability to construct similar meanings in other interactions.

One can imagine this taking place, for instance, in the manner described by the theory of the zone of proximal development (Vygotsky, 1930/1978). Here, a student who is developmentally capable of participating in a certain kind of meaningful interaction with a teacher, parent or older sibling may later internalize the learned ability to engage in that form of meaning making and engage in it with a peer or even internally in his or her mental discourse. While the subsequent internal mental transformations and applications may not be directly visible to an observing analyst, the original learning that took place in the interaction is potentially visible. According to Vygotsky, most learning, especially in young children, takes place socially, interactionally. The hidden, internal learning takes place later, building upon the social experiences. An analyst may want to investigate

interactions among young people or novices, where learning has not yet been internalized as mental cognitive artifacts.

Because instruction consists of forms of interactional practice, it must adhere to the rules of interaction. That is, it must present things in ways that can be seen by participants and whose meaning is made visible to the participants. Because the meanings inherent to instruction must be visibly displayed to the participants, they should be visible to the analyst—under the right conditions. The necessary preconditions are what determine the applicability of the methodology. The conditions can be summed up as: (a) the technical preconditions that determine the adequacy of the video record and (b) the hermeneutic preconditions that determine the analyst's ability to interpret the displayed meanings appropriately.

Video analysis requires (a) the meeting of certain technical preconditions. The video analyst owes his or her existence to the development of digital video technology. Ethnographers and other social scientists have, of course, observed human interactions for a long time, taking notes by hand and more recently using audio recorders. But the interactions within small groups are too complex and subtle to analyze systematically without a more complete record, which one can come back to repeatedly to study. While analog video provided such a record, the real need was only met when one could put the video on a computer and manipulate it frame by frame, zoom in, loop small segments of the sound track, jump around easily to follow lines of inquiry and easily share clips with co-analysts.

Analysis needs a detailed transcript. Depending on the situation under analysis, the transcript may have to include in addition to the words spoken, indications of other sounds, intonations, pauses, gestures, gazes and other non-verbal cues that were visible in the tape. Digital video allows repeated and detailed viewing, as well as the ability to accurately time pauses, in order to produce a useful transcript.

In a situation like a classroom, simply capturing the talk of students with each other during collaborative learning sessions strains the ability of the video analyst even with today's digital equipment. Imagine trying to film the utterances, facial expressions, glances, poses, gestures, inscriptions, computer screens and interactions of a teacher and thirty students in an active, collaborative classroom engaged in an educational innovation. Even if one used hundreds of cameras and microphones and then synchronized the recordings, it would not be humanly possible to follow all that was going on. One must design an interaction setting whose analysis is manageable. By confining the interactions to a sequential stream of messages within small groups in chat rooms, for instance, one not only reduces the volume of data but captures a reasonably complete record of everything that the group of participants shared, already in a textual format.

Video analysis also requires (b) meeting hermeneutic preconditions. A condition for appropriate interpretation is that the analyst has the proper background understanding to know how the participants would interpret the variety of displayed meanings. For instance, do they speak the same language? Assuming that everyone is speaking English, is the jargon of a subculture unfamiliar to the analyst playing a relevant role? Do the students make reference to people or events that the analyst is unaware of? Is there a culture at work in the classroom that the analyst does not understand and cannot figure out from the record?

Even if a whole classroom session was recorded, the analyst may have focused on a few short but interesting episodes and ignored the rest. The question of where to start and stop these analytic episodes is tricky, for they themselves likely refer back to previous episodes and they may be a telling reference for later episodes.

The subjectivity of the interpretation is another important issue. The method responds to this concern by including many points in the analysis where the evolving interpretation is subjected to discussion by groups of analysts, for example in so-called “data sessions” where a dozen or so trained analysts brainstorm about specific episodes and repeatedly view the video clips with detailed transcripts in-hand. Later, when a final analysis is presented at a conference or in a journal, the original data (videos, transcripts, ethnographic notes, etc.—subject to confidentiality constraints) are made available for alternative interpretations. This approach ensures maximal intersubjectivity of the interpretation.

Five Policies from Ethnomethodology

The goal of video analysis is to analyze the practices by which groups of interacting members construct group meaning. Video analysis is founded upon ethnomethodology. Ethnomethodology studies how people (“ethno-”), who are members of communities, construct ways (“method-ology”) of making shared sense of their joint activities. In video analysis, researchers look closely at traces of member activities to study the methods that the members use to achieve meaningful interactions. The meaning-making activities are generally only tacitly understood by the individual members who engage in them, but their meaningfulness is made visible to the group so that it can be shared. Researchers take advantage of this visibility to make the methods explicit. Activities are meaningful in the group perspective. Their meaning is implicitly understood in the individual member perspective and explicitly understood in the video researcher perspective. The phenomenological commitment of ethnomethodology concerns the relationship of the understandings from the different perspectives. Ethnomethodology is a researcher perspective devoted to making explicit the meanings that are understood and taken for granted in the member (individual) perspective and made implicitly visible (for the interacting members as well as for researchers who take the trouble to look) in the group perspective through the utterances, gestures, symbolic artifacts, inscriptions, etc. of the group discourse.

Garfinkel (1967) provided five policies as a starting point for ethnomethodological (EM) studies. These policies are densely worded and complexly interconnected. Therefore, in attempting to summarize them here, I have extracted a key theme from each policy statement and attempted to explain its significance to video analytic research. In particular, I have translated Garfinkel’s terminology (*indifference*, *inspectability*, *relevance*, *accountability* and *indexicality*) into the claim that data for video analysis is *everywhere*, *visible*, *grounded*, *meaningful* and *situated*.

Policy 1: Data Is Everywhere

An indefinitely large domain of appropriate settings can be located if one uses a search policy that *any occasion whatsoever* be examined for the feature that

“choice” among alternatives of sense, of facticity, of objectivity, of cause, of explanation, of communality *of practical actions* is a project of members’ actions. Such a policy provides that inquiries of every imaginable kind, from divination to theoretical physics, claim our interest as socially organized artful practices (Garfinkel, 1967, p. 32).

EM is concerned with the practices people engage in to make sense of each other’s activities. Because human interaction always constructs meaningful order, the EM researcher can analyze almost any interaction (“an indefinitely large domain of settings” of “every imaginable kind”) and discover interesting processes of meaning construction and order negotiation. Groups use meaning-making methods in all social interactions; if one looks closely for these methods they can be found in any domain of interactional data. Of course, the technical and hermeneutic preconditions for analysis must have been met, but that is not a matter of the choice of interactional case.

Sacks (1992) elaborates the argument for being able to discover general methods in most any case of interaction. He argues that for people to be able to understand each other within a complex culture, social practices must be relatively standardized and ubiquitous, and that this has methodological implications for the researcher:

Then it really wouldn’t matter very much what it is you look at—if you look at it carefully enough. And you may well find that you got an enormous generalizability because things are so arranged that you could get them; given that for a member encountering a very limited environment, he has to be able to do that, and things are so arranged as to permit him to. (p. 485)

This means that in order for society to function and for children to be acculturated fast enough to survive in human cultures, people must structure their interpersonal interactions in ways that can be recognized easily. Member methods—despite their vast variety and extreme subtlety—must be ubiquitous and familiar. Consequentially, a researcher can find member methods under any stone, in almost any data set. Conversely, the member methods analyzed in an arbitrary interaction can provide generalizable insights into the structure of member methods in a broad range of situations.

This attributes an important role to case studies. A traditional sociological approach seeks out special events (e.g., examples of best practices) to analyze or imposes laboratory controls on large numbers of cases and computes sophisticated averages. However, the phenomena of everyday practice that are of interest to EM but fall below the radar of other social sciences and conscious folk theories can be studied in depth in arbitrary individual instantiations. Such studies are not “merely anecdotal,” as some critics might suggest. Anecdotal evidence is data based on superficial observations of unscientific observers, often generalized excessively. But EM analyses adhere to rigorous, detailed, intersubjective and inspectable procedures. Furthermore, they only claim to demonstrate how something was achieved in one unique case, although the structure of the methods uncovered may be similar to methods used in many other cases. Case studies are not intended to prove the effectiveness of a specific intervention, but to explore what can, in fact, happen and to investigate the characteristics of actual interactions that are unique but interesting. The criticism that case studies are merely anecdotal is misplaced because it assumes that one is trying to make a universal generalization, whereas a case study is really providing an existence proof that may be

more surprising than a generalization based on common assumptions (e.g., assumptions of which cases are “best practices”).

For instance, an EM analysis of the `SimRocket` transcript would not predict that groups of students under such and such conditions would always learn about paired configurations as a list structure. Rather, it would show how the particular students in that case used methods of repair and explication to establish a shared group meaning, methods that are used in many other interactions. The analysis in chapters 12 and 13 was not intended to conclude whether the `SimRocket` simulation was educationally effective or not. Clearly, the single case could not be generalized to make such a judgment. The case studied was utterly unique. A different group of students might never have engaged in the kind of collaborative discourse that was the focus of the analysis: they might have either seen the list structure immediately or never worked it out. Slight changes in the design of the list (e.g., using a “standard configuration” rather than a “paired configuration”) would have eliminated the problem altogether. If the simulation allowed users to assemble their own rockets (as Chuck in fact proposed), there would have been no list at all to figure out—although the students would eventually have had to construct the equivalent of the list without the help of a list artifact to mediate their work. So, even the smallest generalization would be invalid. Nor could one expect to be able to run multiple trials to average—because each would be a unique experience. But despite this extreme limitation, we were able to discover how a real instance of collaboration actually took place. Our observations of this unique brief moment—despite a variety of shortcomings in the technical and hermeneutic preconditions of our analysis and its very tentative and restricted scope—nevertheless motivated much of the discussion of collaboration in this book.

Through a micro-analysis of a unique case we were able to discover phenomena that permeate collaborative group interaction, but for which our folk theories, intuitions and training did not prepare us. As Sacks (1992, p. 420) said, one can discover from the details of actual empirical cases phenomena that one would not otherwise imagine take place:

A base for using close looking at the world for theorizing about it is that from close looking at the world you can find things that we wouldn't, by imagination, assert were there: One wouldn't know that they were typical, one might not know that they ever happened, and even if one supposed that they did one couldn't say it because the audience wouldn't believe it.

Because any site is as likely as another to reveal the artful practices of rational action, the EM analyst has great latitude in selecting settings in which to do analysis. In particular, any circumstance, situation or activity which participants treat as, for instance, one in which instruction-and-learning is occurring can be investigated for how instruction and learning are being produced by and among participants.

As we will discuss in reference to Policy 3, below, the criteria by which site selection is to be done has to do with how the participants construed what they were doing. The work of the analyst is to conduct an empirical investigation into what participants are doing through their interaction—it is not to impose a theoretical category from outside the interaction. If researchers begin their investigation by seeking out a site that represents “best practice” or “exemplary instruction” or “an example of innovation x,”

they will have begun their investigation by presuming what their investigation is ostensibly designed to investigate. As analysts, we do not presume that we are more informed about learning-and-instruction than the practitioners who do learning-and-instruction. It is not for us to bring to the table preconceived notions or theories of learning and instruction and then see if they are operational within a scene. Instead, our analysis should consist of descriptions of the actions that practitioners perform. These descriptions are specifically oriented to display the sequential organization and orderliness that inform these actions and that these actions are designed to produce.

The analyst does not select data as “cases of x,” but determines what the data is about based on what the data show the participants to be attending to. The researcher’s perspective tries to adopt and explicate the member’s view. As Schegloff (Prevignano & Thibault, 2003) describes the methodology of EM,

The most important consideration, theoretically speaking, is (and ought to be) that whatever seems to animate, to preoccupy, to shape the interaction *for the participants in the interaction* mandates how we do our work, and what work we have to do. (p. 25)

The policy of setting aside or bracketing out externally-supplied characterizations of what participants are doing in conducting an analysis is sometimes described as ethnomethodology’s studied indifference to members’ matters, that is, refusing to impose one’s own interests. It is this indifference that makes ethnomethodological input to a project problematic. Video analysis, conducted under the auspices of Garfinkel’s policies, cannot pass judgment on what might serve as good or bad or even representative practice. EM studies are purely descriptive and cannot be used to form prescriptive judgments. Perhaps these problems can be overcome, however, through clarity about the different perspectives of curricular designers, program evaluators, collaboration researchers and video analysts. EM studies can be used to document, from the research perspective, what members do from their perspective in carrying out educational activities. In so doing, EM studies can produce the data by which designers and evaluators carry out tasks from the design perspective.

Policy 2: Data Is Visible

Members to an organized arrangement are continually engaged in having to decide, recognize, persuade, or make evident the rational, *i.e.*, the coherent, or consistent, or chosen, or planful, or effective, or methodical, or knowledgeable character of such activities of their inquiries as counting, graphing, interrogation, sampling, recording, reporting, planning, decision-making, and the rest. It is not satisfactory to describe how actual investigative procedures, as constituent features of members’ ordinary and organized affairs, are accomplished by members as recognizably rational actions in *actual occasions* of organizational circumstances by saying that members invoke some rule with which to define the coherent or consistent or planful, *i.e.*, rational, character of their actual activities (Garfinkel, 1967, p. 32f).

The idea that social practices are a matter of following culturally defined rules is incoherent, as Wittgenstein (1953) had already argued: Tacit practices and group

negotiations are necessary at some level to put rules into practice, if only because the idea of rules for implementing rules involves an impossible recourse. Although there is certainly order in social interactions of which people are not explicitly aware but that can be uncovered through micro-analysis, this order is an interactive accomplishment of the people participating in the interactions. While the order has aspects of rationality and meaning, it is not the result of simply invoking or complying with a determinate rule. Consider, for instance, the orderliness of traffic flows at stop signs. The smooth functioning in accordance with traffic laws is continuously negotiated with glances, false starts and various signals. Although we do not usually explicitly focus on how this is accomplished unless we take on an analyst's perspective (because explicit awareness is not usually necessary for achieving the practical ends and may actually distract and impede), the signs that are exchanged are necessarily visible to the participants and accordingly accessible to a researcher with appropriate means of data capture.

If we, as analysts, observe rule-like behavior at stop signs, we cannot causally explain this behavior by simply saying there is a social rule that everyone must follow. The members of the group doing the rule-like behavior are continually negotiating what it means to follow the traffic rules in the current context and how they are going to do that. In innovative classrooms, a similar process of rule adoption takes place. If a teacher is given an instructional innovation, she must work out in her situation how she is going to put that innovation into practice in detail (Remillard & Bryans, 2004) and make that visible to her students.

Participants, "as members to an organized arrangement" (Garfinkel, 1967, p. 32), are continuously engaged in the work of making sense or meaning of their own and others' actions. The imputed sense or meaning of an action or of a sequence of actions is not determinate, however, but is instead endlessly open to new interpretation. As Heritage (1984) explained, "The task of fellow-actors ... is necessarily one of *inferring* from a fragment of the other's conduct and its context what the other's project is, or is likely to be" (p. 60). In other words, it is the way that actions unfold that gives them the sense they have. Furthermore, actors are selective in what they treat as relevant so that many aspects of an action's sense remain indeterminate. The only requirement that actors themselves place on their sense making is that it be adequate for the purposes at hand. Meaning, therefore, is "a contingent accomplishment of socially organized practices" (p. 33).

Group interactions are rule-like from the researcher's perspective. But from the member's perspective, the rules are not simply given by social laws that must be obeyed like the physical laws of material objects. A member might take an action that to the video analyst looks like a rule-following response to the situation up to that point. But then it is up to other members to take up new action as part of such a rule or not. For instance, there is a conversational rule that questions should be followed by answers. If someone makes an utterance, the determination of whether that utterance is a question (and therefore part of a question-answer pair) may be made by someone else either providing an answer and thereby establishing the rule, or else laughing and thereby establishing that the utterance was a joke—pending the first person's laughter or objection to their response.

The rule-like behavior is always situated and interpreted within a context of history, activities, artifacts and anticipations. But this context is no more given than the rules that may be followed within it. Members' talk and action has a reflexive character, which is to

say that it is simultaneously “context-shaped” and “context-shaping” (Heritage, 1984, p. 242). While the meaning of any action depends crucially upon the context within which it is performed, the action itself re-shapes the context in ways that will inform the understandability of other actions that follow. This is a mechanism on the micro level of social reproduction, which Giddens (1984a) calls “structuration.”

Rules and context can play important roles in the understanding of interactions from both the members’ and the researchers’ perspective. However, the interpretation of social interaction is a human science and not a physical science (Habermas, 1965/1971), so rules and contextual features are proposed and negotiated within the interaction, rather than being objectively given or analytically proposed. Members may invoke rules as standards within the interactional situation (Pomerantz & Fehr, 1991) to support, justify, rationalize or generally make their behavior meaningful and accountable. Similarly, the discourse may imply that its setting is a certain kind of occasion involving particular categories of participants. Through such interactional moves, members display what social norms and contextual characteristics are salient to their interactions. Researchers should rely on these displays to guide their explicit analyses. If there are warrants in the discourse for interpreting the members as being oriented toward a social rule, then the researcher may bring in a larger understanding of the structures that define that rule but were not made explicit in the discourse (see the discussion of sources of structural Being in chapter 20).

An investigation must rely on the actual practices of the participants as they are engaged in their interactions in order to provide an adequate description of the context of interaction. Such an analysis would constitute a description of the determinate sense of the situation that members construct through their actions. For example, for a researcher to invoke a rule to explain member actions there must be interactional evidence of an orientation to such a rule by the members. In order to document members’ practices in detail, repeated inspectability of these practices is necessary. Video and computer technology provide for this repeated inspectability. This inspectability serves as the only legitimate basis for making claims about such subtle matters as how groups of people took their methods and contexts from moment to moment. As Schegloff (Prevignano & Thibault, 2003, p. 27f, interview of Schegloff in 1996) argued,

These days, only such work as is grounded in tape (video tape where the parties are visually accessible to one another) or other repeatably (and intersubjectively) examinable media can be subjected to serious comparative and competitive analysis. (p. 27f)

In other words, analytical claims about practices must be supported by observable actions of participants, which are evident in the recorded interaction and which establish the facticity and relevance of the claimed matter for the participants themselves. This leads to the recommendation of the remaining three specific research policies.

Policy 3: Data Is Grounded

A leading policy is to refuse serious consideration to the prevailing proposal that efficiency, efficacy, effectiveness, intelligibility, consistency, planfulness, typicality, uniformity, reproducibility of activities—i.e., that rational properties of practical activities—be assessed, recognized, categorized, described by using a rule

or a standard obtained outside actual settings within which such properties are recognized, used, produced, and talked about by settings' members (Garfinkel, 1967, p. 33).

This policy insists on a radical grounded theory approach that derives the categories of the researcher's analysis from the activities of the members. It does not suffice to offer descriptions that depend upon categories defined outside of the situation under study (e.g., student, teacher, gender, learning-disabled, low-achieving, socio-economic status, language ability, etc.) as terms for explaining what participants do or don't do. Garfinkel insists that our theories about member practices must not only be substantiated in the observational data, but should arise from and be grounded in that data. Specifically, we must "bracket out" our pre-existing theories and understandings while constructing our analyses and introducing categories to account for behaviors only when we can empirically demonstrate their "relevance" as evidenced by the talk and activities of the participants. As Schegloff (1991a) observed,

There is still the problem of showing from the details of the talk or other conduct in the materials that we are analyzing that those aspects of the scene are what the parties are oriented to. For that is to show how the parties are embodying for one another the relevancies of the interaction and are thereby producing the social structure. (p. 51)

Further, this policy specifies that actors are not "judgmental dopes" who are incapable of monitoring and acting upon their circumstances. They do not simply follow social laws or rules, but enact these rules (the patterns that appear to researchers as rule-following). They are capable of making choices and they have a shared, if provisional, sense of propriety with respect to what they both can and cannot do and what they should and should not do. While this sense of propriety may or may not be something actors can account for, it is evident in what they do and the way they do it. The work of instruction-and-learning, therefore, as it is actually done, is an ongoing sequence of contingent practices commonly shared among and recognizable by participants. Whether or not a situation is an instance of learning-and-instruction or of successful innovation is not a matter for designers to judge a priori, but for video analysts to demonstrate in their empirical analysis of how the participants took their own activities. This does not mean that it is a matter for the participants to address in post hoc surveys, interviews or focus groups either. For retrospective rationalizations are not the same as the sense making that is enacted in situ. It is up to the video analysis to ground judgments in the traces of the interactive actions of the participants.

Policy 4: Data Is Meaningful

The policy is recommended that any social setting be viewed as self-organizing with respect to the intelligible character of its own appearances as either representations of or as evidences-of-a-social-order. Any setting organizes its activities to make its properties as an organized environment of practical activities detectable, countable, recordable, reportable, tell-a-story-aboutable, analyzable—in short, *accountable* (Garfinkel, 1967, p. 33).

Groups organize their activities in ways that provide for their intelligibility as reportable and inspectable, that is, as meaningful. EM assumes that people ordinarily do things in ways that are inherently designed to make sense. This is a powerful assumption because it allows us to say that actions and the meanings associated with them are sequential in nature and that this sequential organization produces, sustains and is informed by members' shared sense of a local social order. This allows members to recognize prospectively and retrospectively that they are engaged in some specific activity as they engage in it.

When Garfinkel refers to behavior as being *accountable*, the word can be understood in at least two senses. First, members are held responsible for their actions and are accountable to their interlocutors for their utterances and actions; they may legitimately be called upon to provide an explanation or rationale. Second, Garfinkel is contending that all behavior is designed in ways to give an account of the activity as an instance of something or other, i.e., as meaningful. For instance, a group of students might organize their activity to be accountable as a group engaged in doing a class project, in doing a science experiment, in working with a mentor, in being cool, in being teens hanging out. It is the work of the video analyst to document how this making of accountable meaning is accomplished. We will further discuss how social settings organize their own orderly appearance and accountability after reviewing the fifth policy.

Policy 5: Data Is Situated

The demonstrably rational properties of indexical expressions and indexical actions is an ongoing achievement of the organized activities of everyday life (Garfinkel, 1967, p. 34).

Indexical expressions are those whose sense depends crucially upon knowledge of the context within which the expressions were produced. The most obvious examples are expressions that contain deictic terms such as *here, there, I, you, we, now, then*, etc. To make sense of an utterance containing such terms, it will generally be necessary to know who is the speaker, who is the audience, where the speaker and audience are located, when the utterance was produced, etc. Any sentence containing such elements will have different interpretations or meanings depending on the circumstances in which it is produced. Logicians and linguists "have encountered indexical expressions as troublesome sources of resistance to the formal analysis of language and of reasoning practices" (Heritage, 1984, p. 142). Rationalists strive to eliminate indexicality in favor of "objective" propositions; EM acknowledges indexicality's abiding role in situated discourse.

One of Garfinkel's contributions was to note that deictic terms are not the only ones that have indexical properties. Heritage (1984) provides the example of the assessment, "That's a nice one," offered while the speaker and the listener are attending to a particular photograph. What qualifies the picture as nice (e.g., its composition, color rendering, content, etc.) is not made evident by the utterance taken in isolation and must somehow be worked out by the listener by inspecting the object in question. In this way, non-deictic terms such as *nice* are also indexical in use. Similarly, in the SimRocket transcript, when Brent says, "This one's different," each word in this deictic utterance (accompanying a bold, full-body pointing gesture) is itself deictic. As discussed in

chapters 12 and 13, the researcher's attempt to explicate the reference of the term "different" is non-trivial, but highly relevant.

Not only expressions, but also socially-organized actions can have indexical properties. Imagine two people standing face-to-face and one participant reaching out and touching the other. The meaning of this act as a warning, provocation, greeting, demonstration, empathetic gesture, act of belligerence, etc. depends crucially on context, on the nature of the interaction that immediately preceded and immediately follows the touch. (See the concept of "thick description" developed by Austin (1952) and Geertz (1973).)

The fact that the meaning of indexical expressions and actions cannot be determined isolated from the circumstances within which they were produced does not usually present a problem for participants. Brent's indexical exclamation did present problems for his peers, but they managed to resolve this confusion in a few seconds. For starters, participants inhabit the situations within which the expressions and actions are produced and, as a result, are naturally supplied with many resources for resolving their meaning for present purposes. Further, participants have the opportunity to dispel any residual ambiguity through additional sense negotiation. Ultimately, however, all indexical expressions and actions are always contingent and to some degree indeterminate in ways that are deemed acceptable to actors themselves. For Garfinkel, the question of how this indeterminacy is managed in the nonce on a routine basis is at the heart of EM inquiry. It would appear to have similar importance for video-analytic work in the science of computer-mediated collaboration.

The Self-organization of Group Discourse

A central concept in EM is *accountability*. This term defines an important characteristic of group discourse. Although this characteristic is analyzed from the researcher's perspective, it inheres to the group unit of analysis and provides an essential function within the group perspective. It, in effect, makes the group perspective possible.

Let us look more closely at Garfinkel's policy concerned with accountability:

The policy is recommended that any social setting be viewed as self-organizing with respect to the intelligible character of its own appearances as either representations of or as evidences-of-a-social-order. Any setting organizes its activities to make its properties as an organized environment of practical activities detectable, countable, recordable, reportable, tell-a-story-aboutable, analyzable—in short, *accountable* (Garfinkel, 1967, p. 33)

Note that the meaningfulness, sense or accountability of a group activity structure is a function of that setting itself, not a function of people's mental representations about the setting or even of individuals' interpretations of the setting. The setting itself "organizes its activities to make its properties ... accountable." This is not intended as a proclamation about the ontology of reality. Rather, "the policy is recommended that any social setting be *viewed* as self-organizing." That is, it is a methodological principle. In other words, a defining premise of EM is that a researcher should focus on the group unit of analysis and make explicit how the group setting organizes itself. (The view of shared

group reality as self-organizing—as opposed to a view centered on individual minds—will be pursued in chapter 20.)

The pioneers of EM observed that even the most mundane, everyday social settings are organized in ways that seem meaningful to their members, and they posed as a research agenda the working out of the methods of such self-organization. It is an empirical question whether analyses based on this approach are insightful and useful. So far, video analysis and conversation analysis studies seem to offer important views of what takes place in collaborative settings, although their direct aid to design of collaboration support software has yet to be extensively documented.

The relevance of interaction analysis based on EM to CSCL and CSCW has to do with the central role in both fields of meaning making. As expressed in chapter 16, CSCL is supposed to be essentially concerned with the nature of the processes of collaborative meaning making. How do groups make meaning? Garfinkel proposes that *meaning-making* processes consist of the *methods or practices whereby groups make their actions accountable*. This takes place in interaction and discourse. An account of behavior is constructed interactively as people respond to a situation and others take up that response in a particular way, confirming the definition of the context along with an account of the activity. The meaning is constructed not so much by the individual contributions to the discourse as by the ways in which these contributions index, respond to, build upon and take up each other—by the web of interaction.

Let us take two examples from the chapter 12 transcript. First, consider the teacher's utterance, "And you don't have anything like that there?" Initially, the students responded to this as a straight-forward question and supplied an answer in the negative. When this did not elicit a response back from the teacher, they re-constructed the question as a rhetorical question and looked at the list to which this utterance situationally pointed—the list that was on the computer screen, to which the teacher gesturally pointed. From there, Brent started to build an account of how something "like that" was there in the list, by pointing toward a pair of rockets that he saw as satisfying this description. But Brent's statement was just a first step in building an accounting that made sense for the whole group.

Brent's own statement (in turn, recursively) went through a similar process of having an accounting constructed. When he emphatically said, "This one's different," the others at first disagreed. Then gradually they clarified which "one" was being pointed to and how it was different. This involved the shift in conceptualizing comparable pairs of rockets as analyzed in chapter 13. Through this discourse process, the group made Brent's statement accountable. The analysis in chapter 13 from the researcher's perspective made explicit what a full accounting might be like. For the participants, it was enough to say things like Jamie: "compare two n one," Steven: "So I like it how it is" or Chuck: "Oh yeah, I see, I see, I see."

The methods that the students used included a variety of discourse moves: denying, pointing, answering, clarifying, agreeing, completing each other's utterances, repairing divergent references, pausing, interrupting, gesticulating, etc. Each of these fragmentary moves was itself made accountable and only thereby contributed to the larger meaning making. The account of each move and certainly of the larger accomplishment involved an interplay of the discourse context and multiple actions by the discourse participants. In this sense, it was an accomplishment on the group level of analysis.

One can say that the discourse about the list organized itself in order to make itself accountable. It could not have been a successful discourse if it had not done so. The drama captured in the half-minute transcript is the story of how the group discourse organized a story about the list—through the interweaving of contributions from multiple individual perspectives—to the point where Chuck could see the new story, Brent could sit back in his chair relieved that everyone got the story and Jamie could return to the larger story of designing an optimal rocket. One could feel during the long pause preceding Brent’s outburst and the intense student collaboration, while the teacher exercised wait-time, the intense pressure on the group to organize an acceptable story or an accounting of the list to which the teacher directed their attention. The activity of this moment in which the group found itself could not succeed without an effort that managed to achieve an accounting.

The EM notion of accountability provides a plausible and operational notion of group meaning. It is, for one thing, a methodological rather than metaphysical notion. That is, it is not so mysterious and counter-intuitive as the idea of group meaning might appear from the perspective of empiricist folk theories. It can be observed in the concrete analyses of EM practitioners, and judged as to its persuasiveness—in this sense, EM’s notion of group meaning as accountability can be judged based upon the success of its own accountability.

Garfinkel describes a number of characteristics of *accountability*:

1. Accountability is *visible* to members. It is “observable-and-reportable, i.e., available to members as situated practices of looking-and-telling” (Garfinkel, 1967, p. 1). The meaning of social settings is visible to the members of those settings; they can observe and understand that meaning. They can discuss it further themselves and respond appropriately to it. The meaning of a setting is more or less reported, to the extent needed for the practical purposes of the discourse interaction. The meaning-making process as a set of member methods is generally taken for granted and not reported in the group discourse. However, traces of those methods and how they have been taken by the group are observable to researchers, who can make them explicit. Thus, group discourse by its nature makes group meaning visible—for both members and researchers, in their own ways.
2. Accountability is an *accomplishment* of groups. The meaning is not something distinct and separable from the social settings, activity structure or group context. “Their rational features *consist* of what members do with, what they ‘make of’ the accounts in the socially organized action occasions of their use” (p. 4). The accountability is an accomplishment of the on-going interaction of the occasion; it is an emergent feature of the occasion itself.
3. Accountability is *indexical*. Where other sciences try to formulate abstract generalizations, EM insists that meanings of practical group interactions are necessarily tied to concrete contexts that they reflexively specify and that they index. One can try to substitute objective terms for deictic references, but this process is in principle incompleteable. The scientific attempt to render every description in “objective,” quantifiable, classifiable, generalized categories by substituting explicit terms for deictic ones “remains programmatic in every particular case and in every actual occasion” (p. 6).

4. Accountability is *reflexive*. Discourse takes place on multiple levels simultaneously. Or, discourse can be interpreted in multiple, mutually consistent ways. When a group discusses some content, they are also at work in their discourse making their discourse accountable. For instance, the discussion about comparing rockets was also a discussion about repairing misaligned references and constructing a story about how to analyze the list. The students focused their comments on the rocket content—which numbered rockets had which attributes. The research analysis, by contrast, focused on the meaning-making process. In general, members are little concerned at an explicit level with the account that they are creating—unlike the researchers, who are not much interested in the discourse content except as it reveals the accountability. The reflexivity of accountability has to do with the fact that the two levels are part of a single reflexive process: “members’ accounts ... are constituent features of the settings they make observable” (p. 8).
5. Accountability is *tacit*. This is related to the fact that members are not much interested in the methods they use for making their discourse accountable. One can say that the methods of accountability are themselves not accountable. Although member activities accomplish meaning making on multiple levels, “for the member the organizational hows of these accomplishments are unproblematic, are known vaguely, and are known only in the doing which is done skillfully, reliably, uniformly, with enormous standardization and as an unaccountable matter” (p. 10).
6. Accountability is *shared*. The EM notion of accountability sheds light on the discussion in chapter 17 of shared meaning, common ground and group cognition. Accountability can be seen as the establishment of a group meaning by reference to rule-like methods. Shared agreement is then seen to be an interactive accomplishment in which a group establishes that the discourse is to be accounted for in terms of a specific method or rule. “To see the ‘sense’ of what is said is to accord to what was said its character ‘as a rule.’ ‘Shared agreement’ refers to various social methods for accomplishing the member’s recognition that something was said-according-to-a-rule and not the demonstrable matching of substantive matters. The appropriate image of a common understanding is therefore an operation rather than a common intersection of overlapping sets” (p. 30). Here, the agreement that is so fundamental to social interaction, collaboration and intersubjectivity is clearly not viewed as a matter of overlap among sets of mental representations in members’ minds—as the common ground approach seemed to conceive it—but as the successful achievement of accountability of a group discourse.

The EM approach, with its central notion of accountability provides a plausible way of thinking about how the self-organization of group discourse provides a basis for making group meaning visible.