

# A Study of the Foundations of Artifact-Mediated Collaboration

Nathan Dwyer, Daniel D. Suthers

Laboratory for Interactive Learning Technologies  
Department of Information and Computer Sciences  
University of Hawai'i  
ndwyer@hawaii.edu, suthers@hawaii.edu

**Abstract.** The premise of this work is that, like language, the meanings of written representations are contextual and their affordances are appropriated in sometimes-unexpected ways. This situation presents a dilemma for designers of collaborative learning technologies: there is a need for representational and interactional tools that guide and support learning through cognitive and social activities, but predefined mappings between interface elements and functionalities may be too rigid. The study reported in this paper attempts to address this dilemma by identifying the strategies that people communicating via flexible written representations use to manage their interaction, and how they appropriate the affordances of media to carry out these strategies. We analyzed how people appropriated paper-based tools for collaboration under conditions approximating online interaction. Regularities were observed in the use of limited but polymorphic repertoires for communication and expression of attitude, functional and coordinative use of space, the presence of simultaneous threads, and strategies for interruption and context setting. The results suggest a new generation of collaborative technologies that include support for multi-faceted and parallel interactions, lightweight tools for expressing attitude, context representations, and scaffolding for automatically detecting and supporting emerging conventions.

**Keywords:** Descriptive Studies, Video Analysis, Interactional Practices, Representational Affordances, Shared Workspaces

## INTRODUCTION

Studies of spoken language have established that the meanings of utterances are contextual and negotiated only to the level of agreement needed to support action. The premise of this work is that the same is true of nonlinguistic representations. Meanings of representations are not fixed in advance, but change according to context. Efforts to provide users with a visual language for reasoning or argumentation have often encountered difficulty creating a functional notation (van Bruggen & Kirschner, 2003) and in getting people to adopt such a notation (Conklin, 2003; Selvin, 2003). Also, users resist the very idea of categorizing their thinking, although they do seek representational aids in organizing ideas (Shipman & McCall, 1994). People make flexible use of representations, and the perceived affordances (Norman, 1988) of representations are appropriated in sometimes-unexpected ways. For example, Dillenbourg & Traum (1999) had participants use synchronous chat and an electronic white-board in a MOO environment while solving a murder mystery. They expected that the white-board would be used for disambiguating spatial references through its two dimensionality and drawing affordances. Instead, the most important affordance of the white-board for participants turned out to be its persistence: information that had to be recorded permanently was written in the white-board.

This situation presents a dilemma for designers of computer supported collaborative learning (CSCL) or collaborative work (CSCW) systems, or indeed computer-mediated communication (CMC) in general. (We will refer to these collectively as collaborative technologies.) There is substantial work on replicating the properties of face-to-face (FTF) communication by using high bandwidth video and audio, and techniques such as clever placement of cameras and screens for accurate conveyance of gesture and gaze (e.g., Kato *et al.*, 2001). Yet, others believe that collaborative technologies offer unique opportunities (Dillenbourg, in press) and should go "beyond being there" (Hollan & Stornetta, 1992) by exploiting the special properties of computational media in ways that make distance interaction more effective. This message resonates with educators' recognition of the need to guide and scaffold learning. Whether the application is intended for learning or work, it is not enough to simply attempt to provide a channel of communication as rich as face-to-face communication. We truly realize the potential of these technologies only if we use them to guide and enable more effective learning and problem solving practices. This point applies equally well to face-to-face interaction. If the richness of FTF were sufficient to solve problems of learning and collaboration, we would not see interest in technologies that support these activities specifically in FTF contexts.

The dilemma manifests when designers try to build representational and interactional tools that guide and support cognitive and social activities. User interfaces tend to define rigid mappings between GUI elements and functionality. Software tools that are easy to build with typical GUI toolkits are not a good match to the flexible

nature of human communication, nor do they adapt to the changing needs of the user. Constrained representational tools may guide learners in their enculturation to a new field, but must be kept simple for this learning period and this simplicity may soon become too constraining. Conversely, a set of tools that is sufficiently complex for supporting experts in a field can be daunting, and dissuade novices

The present study attempts to address this dilemma by identifying how people appropriate flexible representations to meet their needs—essentially, to identify what Garfinkel calls “member's methods” (Garfinkel, 2002) for synchronous collaboration via written representations—so that we can build CMC tools with affordances that support the kinds of flexibilities observed. Like Sacks, we are “trying to find the machinery” (Sacks, 1984) of a kind of social interaction. Whereas Sacks focused on conversation in naturally occurring interactions, we are focusing on the mechanisms of intentional collaboration via written media. Despite these differences, we claim that it is valid to treat these interactions as locally-managed and socially constructed. We want to answer such questions as: When people communicate via written means, what strategies do they use to manage the interaction? How do they appropriate the affordances of media to carry out these strategies? How can our CMC tools provide those affordances while also offering task-specific guidance? This paper addresses the first two questions empirically and discusses implications for the third.

We cannot adequately answer such questions with studies of CMC, nor of FTF interaction alone. A third strategy is needed and taken by this study. We cannot effectively conduct the study with an existing CMC technology since any CMC technology we chose would carry with it the very assumptions of collaborative technology design that we are questioning. Exploratory development of collaboration technology is another alternative, but comes with a high cost of implementing each iteration of the software. It would also be a mistake to conduct the study with unrestricted FTF interaction because there are too many differences between FTF and communication technologies to create a reasonable mapping from one to the other. FTF interaction includes many subtle cues that are difficult to replicate online, and we want to find the special advantages of artifact-mediated communications that might not be evident in FTF interaction. Therefore we take a middle road. Our strategy is to start with FTF, but restrict or remove some of its features that are especially hard to replicate online and add other features such as persistence that are advantageous online. We require that people communicate with written representations, but do so using very familiar and flexible tools -- paper office supplies—so we can get a sense of which affordances of flexible representational tools participants take up, and for what communicative functions. Many attributes of the tools were varied so as to highlight invariance in communicative functions across different permutations. In this paper we report on the range of ways in which the tools were exploited and on the mapping between actions on tools and communicative functions, and discuss implications for design.

## **METHOD**

The purpose of this study was to discover strategies or methods that people use to collaborate through shared written representations. This understanding is sought both independently of and in reference to the representations used: We want to know what kinds of communicative or coordinative functions people consistently attempt to implement independent of the representation used, and we want to know how people appropriate the affordances of specific representations for these purposes. Our methodological strategy is to vary the representational tools provided and look for invariants across the different situations. This differs from an experimental design, which attempts to control as many factors as possible and show there is a difference between experimental groups correlated with the one thing that varies between those groups. We are not making comparisons between experimental groups, and the variation in participants and materials between the sessions is desirable, as it strengthens claims of generality. Our method also differs from the micro-analysis of typical Conversation Analysis. This study focuses on identifying interactionally constructed mechanisms. Understanding the details of how the mechanisms are actually constructed is left for later work.

### **Participants**

The study involved six pairs of friends recruited from community college and university students. Their ages ranged from 18-24, with an average of 20.9. They had a diverse set of majors, including Digital Art, Nursing, Environmental Studies, and Computer Science. Every pair except one had known each other for multiple years and all had consistent social contact. An additional six pairs were recruited for pilot studies. Unless noted otherwise, the observations below do not include data from the pilot studies.

## Environment

A large table was prepared with a screen suspended above it such that participants seated on opposite sides of the table could not see each others' faces (see Figure 1). The seating was arranged such that participants could reach and therefore manipulate the entire workspace.

A video camera was placed above and to the side of the participants, positioned in the same plane as the screen to minimize the extent to which the screen blocked the camera's view of the workspace. The camera captured both participants and the entire workspace, except for a thin line blocked by the screen. Digital video output was streamed to a hard drive in real time.

## Materials

In all cases, the table was covered with a secured sheet of butcher paper. All pairs had access to tape, rulers, scissors, string, paperclips, and sticky labels, dots and stars, and all pairs were provided with the same collection of various writing utensils, including red, green, blue, and black white board markers and ball-point pens, and blue and black permanent markers.

Different pairs were provided with different kinds of paper products. Three pairs of participants were given unrestricted amounts of office supplies such as varying sizes and colors of paper, multiple sizes of index cards, and multiple sizes and colors of Post-It(tm) (sticky) notes. One pair was only given 3x5 inch index cards. Two pairs were given a single, large sheet of 2x3 foot unlined paper. The original intent was to have two pairs for each condition, but after the first study session using only 3x5 cards, the data from the unrestricted office supplies appeared more promising.

Office supplies were chosen partially for their familiarity. It was assumed that participants would not need time to learn their affordances. Also, office supplies have a demonstrated history of versatility and effectiveness. The three sets of materials were chosen for the specific affordances they provided. Individual pieces of paper and index cards can be moved around and repositioned in relation to each other, and Post-It(tm) notes can be attached and detached from other materials. The index card only condition removes the attachment affordance, while the large sheet of paper disallows repositioning the participant's contributions.

## Procedure

Pairs of participants were given discussion topics, and asked to brainstorm ideas, discuss them, and come to some kind of final agreement. The pairs were given three topics chosen from a pool of five "wicked" (Rittel & Webber, 1973) problems, such as: How do we preserve Hawai'i's environment? Space aliens are coming; how should we respond when they arrive? What is the appropriate relationship between science and religion? Wicked problems are typified by the lack of clear evaluation metrics for any answer as well as the lack of a well-specified process for approaching them. This required the participants to collaboratively develop processes for collaboration and evaluation, and to negotiate when they had come to the end of the interaction. The order and selection of topics assigned to the pairs were permuted to create the greatest variability.

Every effort was made to support the greatest amount of flexibility in the representational medium. On the other hand, several restrictions were placed on the participants in order to approximate limitations of online communication. Since current online communication is predominantly text-based, we chose to limit visual and verbal channels. Participants were required to communicate entirely through the use of the pens and materials, and because of the screen, participants were unable to communicate using facial expressions. They were also asked to remain silent while working on the assigned problems. The participants' hands and arms were visible to each other. Given that deixis is so fundamental to communication, and since tele-pointers and avatars are viable CMC tools, this seemed to be a reasonable allowance.

Problems were printed on sheets of paper and given to the pairs one at a time. Pairs were allowed to ask for clarification about the problems before they began each session. They were told that they should collaboratively analyze the problem in as much detail as possible, and that they needed to come to a final conclusion they both agreed on. Pairs were given 30 minutes to work on each problem. The experimenter kept notes on his observations during the study sessions. After each problem the pairs were interviewed on their conclusion, what they thought of the interaction, difficulties or issues with the procedure, and any other reaction to the session. The experimenter also used this time to verify his interpretations of the session activities.

## Analysis

The video data was analyzed in multiple passes. We approached the video with no predefined theory or system, and made every effort to let the data guide the analysis. A custom application was developed to support variable



Figure 1: Study Environment

speed review of the video and a simple annotation system. Initial review of the video was done at normal speed, to reinforce awareness of the contents, and then at double or triple speed to get a sense of the larger-scale recurring patterns. Initial reviews of the video identified several of the most obvious interaction patterns—generally those that dealt with use of the space and interaction structure. Successive reviews looked specifically for these patterns and identified several more. The entire collection of video (approximately 8 hours) has been reviewed multiple times, and several interesting segments have been studied in more detail.

## OBSERVATIONS

With the exception of one pair, all the participants reported that while it was constraining to communicate with only the written word, they still felt they were able to communicate successfully. One participant in the pilot study reported that he actually preferred the written format in some ways because it was harder to be interrupted by the other participant. The S5\* pair was an exception. This pair evidenced a high degree of interpersonal conflict, and didn't come to agreement on any of their three topics. The observations reported below exclude the S5 pair. A summary of their interactions is provided later.

In this analysis, we were specifically looking for patterns related to the mechanics of collaboration, and not the content of the interaction nor the practice of sense-making (which may be considered in future analyses). After some general comments about the structure of the interactions, we describe six categories of specific mechanisms that we identified: Functional Spaces, Simultaneous Threads, Placement, Interruption, Setting Context, and Repertoire.

### Structure of Interactions

The interactions all resembled informal conversations. Topics and ideas introduced by each participant were explored only to the point where the two participants agreed on either the topic's relevance or a general conclusion. Few topics were revisited, and usually only to verify agreement before the information was incorporated into the final product.

The content displayed a topically episodic structure. The participants moved back and forth between simultaneous individual work and collaboration. Collaborative interactions were usually initiated by one participant, and consisted of contiguous contributions (possibly in parallel with other activity) until the topic was resolved. The resulting artifacts reflect this structure. On a single large sheet of paper, areas can be identified for each topic addressed. Pairs who used 8.5x11 pages generally confined each topic to a single page, and pairs who relied on smaller materials almost always introduced a new post-it or 3x5 card to start a new topic.

There was no evidence of complex information management. For example, the physical nature of the materials was not exploited to investigate connections between ideas or to propose categorical groupings or inclusions.

### Functional Spaces

Every pair used some mechanism to separate the text of the argumentation from the text of the conclusions. Once these spaces were designated, participants resisted altering their purpose. In several cases the pair continued to squeeze the argumentation text into a constrained area long after this stopped being possible without interfering with the prior record. One pair in the single large sheet of paper condition split the paper into a “Work Area” (~30%) and a “Plan” (~70%) area (see Figure 2). When the “Work Area” had been completely filled with text, the “Plan” area was still almost completely unused. Rather than re-designate space from the “Plan” area, the participants used larger markers to write over the top of the existing text in the “Work Area”. This behavior was not unique. Many pairs constrained their contributions to an impractical writing area while leaving large areas designated for results unused.

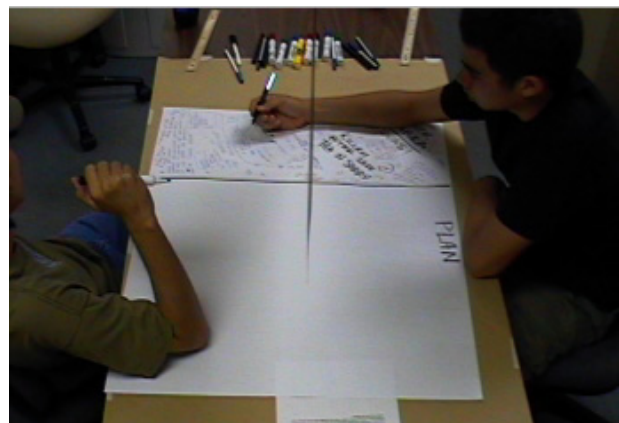


Figure 2: Functional Spaces—the work area is full, while the area for conclusions remains unused.

\* The sessions and participants are identified by codes. S5 identifies the fifth pair to participate. S5L and S5R respectively identify the participants sitting on the left side and right sides of the table in relation to the camera.



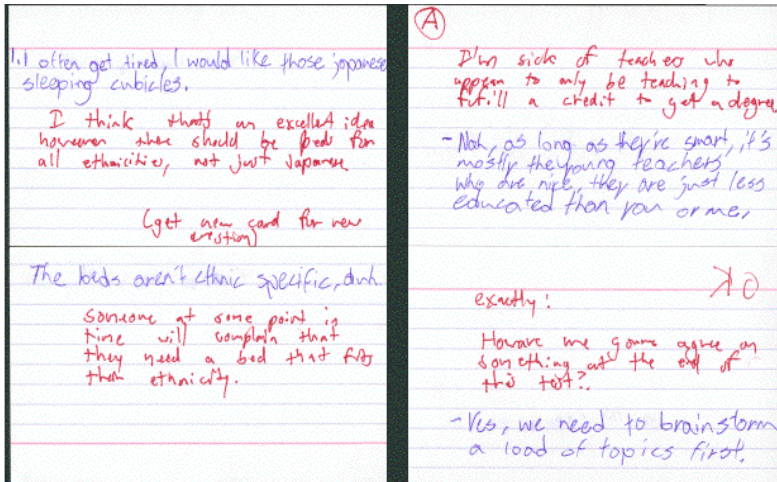


Figure 3: Simultaneous Threads—The participants used different pen colors. These cards were exchanged repeatedly.

Figure 3). This had the added benefit of indicating when each person was done writing, and produced a spatially coherent artifact oriented to the topic. Participants using a single large sheet of paper tended to group contributions to a thread, but occasionally participants in this condition carried on synchronous threads by alternately writing on the paper as close to their partner as possible (i.e. on the far side of the table from themselves). This resulted in several topic threads that were not at all spatially coherent, but this didn't appear to interfere with the interactional coherence of the exchange. More often, though, topics were confined to a specific area on the paper, and there are easily observable spaces between the areas for each topic.

Pair S3 was an exception. Each maintained their own material to write on, and would alternately write and then read what their partner had written. They didn't reorient their material so the other could read it—they consistently read the other's text upside down. The S3 pair consisted of members of the deaf community, and while this use of the materials resembled a common practice in that culture, the participants reported that the usage was not a conscious choice. This use of the materials produced interactions consisting of a single thread, with each contribution alternating between separate pieces of paper.

Pairs using mobile materials tended to keep related information on a single piece of paper. When a participant made a contribution to an existing thread, he/she did so by adding the contribution to the piece of paper that contained that thread. On the single large sheet of paper, text ran in all directions, and while related contributions were usually added near the related text, there was occasionally no space to add a contribution. In this condition, pairs made use of lines and arrows to connect their contribution to previous text. These connecting arrows were sometimes used even when the contributions were in proximity to each other or the connection was obvious.

Pairs displayed no cognitive difficulty in managing multiple topics. The simultaneous topics on the different 3x5 cards remained coherent, and the participants displayed no difficulty keeping track of conversational threads. On the large sheet of paper, participants had no difficulty being interrupted during an individual contribution to focus on their partner's topic. The interrupted party went back to his/her individual work with no problem.

## Placement

The pairs using multiple materials made remarkably consistent use of position and placement. The tabletop workspace allowed objects to be moved around and overlapped. Unsurprisingly, every pair in these conditions consistently contributed to the current topic using the space directly between the two of them. Objects that were no longer current were set to the side.

More interestingly, participants positioned objects between themselves so as to mediate access. Individual contributions were consistently made with the object close to the writer. A standard pattern was for the writer to position a card close to his or herself, write a contribution, and then move the



Figure 4: Placement—the participant on the right is adding to the conclusions, while his partner is starting a new topic.

## Simultaneous Threads

The nature of the materials afforded simultaneous contributions from each participant. The large sheet of paper provided ample space for both participants to write concurrently. Smaller, individual pieces of paper allowed each participant to have their own writing area.

In all but one pair, there were frequently two concurrent threads of contribution. When participants used materials that could be positioned (i.e. sheets of paper, 3x5 cards, etc.) each piece of paper usually represented a single topic thread. The pair would alternately write on the paper and then pass it across the table to their partner (see

card to the other participant's side of the table. The position also regulated the contributions of the writer's partner. When a material was in the center of the table, both participants would regularly write words and complete sentences on it. When the material was close to one participant, however, the other participant limited their contributions to pointing or tapping or writing individual symbols, most often a question mark.

The more collaborative interactions took place nearer the center of the table (see Figure 4). When the two participants were engaged in negotiating an agreement, the interaction often took place on a single piece of paper positioned in the center of the table. The shared paper was turned repeatedly so each participant could write easily, but participants tended to orient the paper sideways, allowing each participant the same ability to read and contribute, and rarely moved the paper closer to their own side of the table. Papers used to record conclusions or final results also often stayed near the center of the table and usually remained oriented sideways. This sideways orientation of the shared conclusion was also used in one of the large sheet of paper conditions.

## **Interruption**

All the pairs used a similar protocol for interruption. To get the other's attention, a participant would touch the other person on the arm, or tap repeatedly on some text. These interruptions were often not responded to immediately. Participants were able to queue the request for attention until they were done with their immediate task. This usually meant that the person would finish writing a sentence or conclude a thought. Usually, the person making the interruption did not insist on an immediate response from their partner. After a participant had requested their partner's attention, they would often go back to work until the partner finished their task and changed focus.

## **Setting Context**

When the members of a pair were working individually, and one of them wanted to engage the other, two specific behaviors were used for setting the current topic.

When a new topic was introduced, a participant would usually write an initial sentence and then bring it to the attention of the other participant before continuing. With materials that afforded positioning, this generally involved the writer moving the material to their partner's side of the table. Several mechanisms were used to negotiate when materials could be exchanged or moved. Either the participants would wait until they were both done writing and then swap materials, or the writer would place the material on their partner's side of the table, sometimes intentionally obscuring the partner's current work. With a single sheet of paper, the writer would produce the introductory sentence, get their partner's attention, and then tap the new text. This was often followed by the writer running a finger under the text, indicating what the partner should read.

When a participant wanted to return to a previous topic, the order of actions was reversed. The participant would indicate previously written text by tapping on it or underlining it, and then write a related sentence. This related sentence sometimes did not need to be in proximity of previously written text. Participants had no problem understanding that the underline was meant as an indicator, and not for emphasis. When the participants re-read the text later, they did not appear confused by the underlines or other markings that had accrued.

## **Repertoire**

Despite the availability of a wide variety of materials, every pair constrained themselves to a very limited subset. Even when the widest variety of materials was made available, all pairs tended to use only one or two materials almost exclusively. New types of materials were generally introduced to distinguish types of information, i.e. in a session recorded entirely on 3x5 cards a pink post-it was used to record conclusions.

In addition to limiting their choice of materials, all the participants made use of a remarkably limited set of gestures, deixis, and symbols. Women were considerably more likely to use hand gestures than men. Still, these gestures were mostly limited to an approval gesture (e.g. thumbs up) and a questioning gesture (e.g. hands spread palm up, like what usually accompanies a shrug). Actions that related to the artifact consisted of a variety of pointing gestures. Participants pointed at, tapped on, or ran a finger along (under) artifact elements. Gestures involving both hands were apparent, e.g. indicating some written text with one hand, either pointing, tapping, or underlining, and then tapping the area being used to record conclusions with the other hand.

A similarly limited number of symbols were employed in the written artifacts. Symbols were almost never used independently, but almost always as an annotation or in reference to some other piece of text. The symbols used regularly were question marks (by far the most prevalent), arrows, smiles, stars, and check marks. Two different pairs of males each drew representations of a hand with the middle finger extended, but only once each. Despite a variety of labels and stickers, pairs only made use of stars and colored dots.

Actions, gestures, and symbols were polymorphic. Pointing, for example, might be used to indicate suggested topics, related information, reminders, request for clarification, or illegible handwriting. Participants used underlining to indicate emphasis, to relate the current discussion back to previous discussion or the problem statement, to set the context of a subsequent contribution, and to indicate repetition. The meaning of question marks and other symbols were similarly context-sensitive.

## **Conflict**

The sessions were remarkably consistent in the high degree of collaboration, the success of arriving and shared conclusions, and the use of similar mechanisms. The S5 pair was an exception. They displayed a high level of hostility towards each other, and they failed to arrive at a conclusion to any of their three topics—the only pair to fail to do so.

The two women had very different demeanors. S5L was very aggressive while S5R tried to be more conciliatory. S5L tended to make long, individually produced contributions, but then gave short, negative responses such as “no” or “who cares?” with a considerable amount of underlining. The most interesting aspect of the pair’s interaction, however, is that the communicative methods they used were consistently the reverse of the other pairs.

Rather than the episodic structure seen in the other pairs, the S5 pair would pursue a topic only until there was a conflict at which point S5L would usually turn the material over or introduce another piece of paper and start a new topic. S5L tended to ignore her partner’s interruption requests, but was insistent that her own be attended to immediately. The pair never settled on a specific place for conclusions. S5R tried to designate such a space, but S5L did not adhere to that designation. In one session, S5R attempted to maintain interaction on a centralized sheet of paper, but when S5L contributed, she would pull the paper to her side of the workspace and halt the interaction by writing long tracts of text. During another session, S5L would often reach across the table and write whole sentences. One entire segment takes place on S5R’s side of the table.

It is interesting that the only pair that did not display many of the coordinating mechanisms just discussed also displayed a high level of conflict and lack of success. In this sense, S5 is the exception that proves the rule.

## **DISCUSSION**

Despite all the variation introduced into the study sessions, there is considerable evidence of consistent communicative needs and methods for meeting these needs. This consistency spans differences in topics, materials, and individuals. That these methods were consistently recreated indicates something about the importance of these needs. Implications of addressing these needs in a CMC environment will be discussed further below.

Some environments provided affordances that naturally mapped to a specific task, e.g. relating two ideas by placing them in proximity with each other. In the conditions where the materials did not provide these natural affordances, participants did their best to approximate them. For example, on the single large sheet of paper arrows were used to approximate collocating related ideas. While pairs with multiple materials tended to reserve specific papers for their results, in the single large sheet of paper condition people approximated this by marking off areas reserved for conclusions. The methods participants developed were certainly driven by the affordances of the materials they were using. However, the fact that pairs given less appropriate materials approximated methods by appropriating the available affordances implies that they weren’t just responding to the environment as it was given to them, but that they had some idea of specific communicative methods that they felt were required.

These observations might lead one to the conclusion that the specifics of the environment make little difference. It is possible that given any environment, people will create tools appropriate to the environment, adapting to the limitations of their tools and “making do” with whatever affordances are available in order to perform important communicative tasks. Conversely, it might be argued that we could determine the “best” affordances, and simply ensure that any collaborative environment supported them. However, the effectiveness of an affordance cannot be judged in isolation. Different affordances and groups of affordances will map more “naturally” to different activities. A comparative approach should be taken to determine how collections of affordances more effectively support any specific collaboration method.

### **Support for Complex Interaction**

#### *Flexible Topical Structure*

The pervasive use of multiple, simultaneous topics threads mirrors the kinds of interaction being studied in multi-user text chat environments (Herring, 1999; O’Neill & Martin, 2003) O’Neill and Martin make the case that temporal proximity is not as important as interactional coherence, and that participants in a text chat environment will develop mechanisms for repairing misunderstanding. In fact, they found that rather than being limited by the lack of temporal proximity, participants incorporated it into their communication as a source of humor. Like people using text chat, participants in our study were able to maintain interactional coherence via the affordance of persistence rather than strict alternation of contributions. Similarly, the text of contributions was not constrained by direction, succession, or strict alternation. Participants made full use of the ability to add contributions with minimal regard to format despite the fact that this produced messy, complex workspaces. These kinds of expressive freedom should be explored in the next generation of collaboration tools.

### *Managing Multiple Topics with Interaction Contexts*

The overall structure of the interaction followed a sort of stream-of-consciousness series of topics. There was no advance organization or agenda setting. Instead, the participants responded to each other's contributions as they were brought into the workspace. Rather than imposing a mechanism for ordering and managing topics, the participants managed their focus on each topic through the affordances of the physical workspace.

Participants seemed to move the various topic threads between several states. Participants used position and orientation to invite a response, to initiate collaboration, or to indicate ownership (and the corresponding limitation on their partner's contributions). This happened with multiple topics at a time, indicating a model of discourse that is less linear and instead relies on opportunistically advancing elements of the overall interaction. Collaborative software could support this model by providing better management of multiple topics including simultaneous awareness of the state of each. These observations also suggest a collection of loosely defined "soft" workspaces, each representing a context of interaction that includes topic information as well as affective information and the current participant expectations. The software to support this would have to support a simple mechanism for switching attention from one topic to another as well as merging and branching the contexts.

### *Prioritizing Contributions*

Not all contributions are of equal importance. The use of interruption indicates that participants want to bring some contributions to their partner's attention immediately, while others can wait until the partner is available. Within the confines of the synchronous environment, there seems to be some variability in just how synchronous an interaction needs to be. Collaborative software could support this management of prioritization by providing mechanisms for passive contributions to a shared artifact along with more proactive tools that alert participants to certain contributions as they are being made.

### **Ease of Indicating Attitude**

The majority of contributions to any session fell into two categories: contributing new information and indicating an affective response to that information. Topics were addressed just until both participants had indicated (dis)agreement or (dis)approval, not until they had been exhaustively examined. Each participant used the ability to quickly express an affective response as a recurring tool to help shape and direct the interaction. It is notable that participants indicated attitudes with symbols and hand gestures rather than written out sentences. The ability to easily and continuously express attitude is fundamental to conversation, but generally absent from current collaboration technologies. Even the use of emoticons and similar tools requires a conscious mapping of response to representation.

For the participants, the affective layer of information took the place of thorough data analysis. In the artifacts there is a wide variety of content types, but there is a notable absence of any sort of organizing mechanisms. Aside from defining work and conclusion spaces, there are almost no explicit rating systems, pro and con lists, or inter-relation between the topics. Since indications of approval and agreement were often made with non-persistent actions, the participants were required to remember the status of the various topics and of the interaction as a whole. Without a visible record, both participants would have had to assume that they were each remembering things in the same way. There's an upper limit on the complexity of data managed this way. After a certain point, it becomes necessary to impose some organization so the information can continue to be used effectively. Although collaboration technologies can help with this, fixed organizational schemes would carry with them the problems of inflexibility that motivated this study. We should explore collaboration technologies that automatically infer the implicit organization from affective contributions and artifact manipulations, and facilitates the application of this organization as the need becomes apparent.

This discussion suggests two important limitations of this study; the limited time frame and the dependence on personal opinion and knowledge (rather than external data and/or formal evaluation criteria). It is possible that a longer time frame or the requirement to use external information and formal criteria would have motivated participants to invent more organizational mechanisms. Also, writing all contributions by hand limited the amount of text that could reasonably be generated in 30 minutes. This seemed to dampen enthusiasm for extended debates or long explanations.

### **Simpler Tools, Emergent Conventions**

The participants in this study only developed a small number of communicative methods, and once a participant pair chose a material or method, they tended to use it for the duration of the session. Even when the method was found to be inefficient, or they agreed on its limitations, they fell to using it again for the next topic. Similar self-constraint on the use of available materials and tools has been documented in usage studies of complex software such as word processors and spreadsheets. (Nilsen *et al.*, 1993) notes that users only learn a subset of the application's functionality—usually just enough to accomplish their normal tasks. It is possible that fewer yet more flexible tools in collaborative software might encourage users to make more creative and personalized use of the environment.



There is a similar self-imposed limit on the use of notational conventions. In the few cases where conventions were attempted, they either quickly became confusing, were immediately ignored or miscommunicated between the participants. The effort of remembering and applying a convention outweighed its perceived value. Collaborative technologies should be designed to reinforce the use of notations by allowing conventions to emerge from the interactions without them being explicitly defined.

## **User-Driven, Complex Actions**

### *Polymorphic Actions*

In contrast to the limited number of actions, the communicative functions performed were numerous. Every action was used for multiple purposes. Actions such as pointing, underlining, and drawing an arrow do not by themselves have concrete semantic content or indicate a specific communicative function. Each of these takes on meaning given the context in which it is used. Each of these actions could be combined and overlapped with other actions as well as modified, giving rise to the wide range of meanings we observed. For example, tapping on a piece of text that one has just written draws attention to the new contribution, while tapping on text just written by one's partner may indicate agreement, or if followed by further writing may indicate the intent to comment. These combinations were additionally modified based on the interactional context in which they occurred. The position or ownership of the indicated text changes the meaning of the writing that followed. Finally, the nature of these composite actions was additionally modified by the perceptible attitude of the person who was performing them. The speed or intensity with which the actions were performed further extended the already broad palette of possible communicative acts. For example, a participant may have tapped more gently to show agreement or more energetically to show insistence.

Despite polymorphism, the intention of each action was almost never misconstrued. Rather than create any formal definitions, participants were willing to expend the mental effort to perform continuous (re)interpretation of these actions and their context. For example, post-interviews with the S6 participants revealed that even when there was an awareness of a mismatch in interpretation (P1 knew that P2 had misunderstood the meaning of a notation used by P1), the participants preferred to revise their understanding of each other's behavior rather than produce external documentation (P1 interpreted P2's actions in light of the misinterpretation).

The number of possible meanings accessible to the participants could not be reasonably represented using the typical "single tool = single meaning" paradigm available in existing user interface development toolkits. We should investigate the creation of simple, flexible tools that support creative combination and repurposing. Allowing users' own repertoires of symbols and actions to emerge from their interaction will give rise to richer and more actor appropriate tool sets.

## **CONCLUSIONS**

This study was conducted in an environment significantly different from either typical conversation or typical on-line communication. Face-to-face conversation is generally not persistent, and relies heavily on a wide range of non-verbal cues. Typical online environments provide a structured set of tools to facilitate communication, but don't support complex gesturing or physical interaction. This study environment attempted to marry the limited communication channels of the online environment with the flexible representational abilities of pen and paper. Participants' level of engagement suggests that this marriage was successful.

The data gathered from this study shows a great deal of consistency at the structural level of artifact-mediated communication. Analyzing interaction from the bottom-up gives a sense of how people act "naturally" and suggests several possible implications for the design of collaborative systems. A new generation of collaborative technologies could include support for multi-faceted and parallel interactions, lightweight tools for expressing attitude, context representation, and scaffolding for automatically detecting and supporting emerging conventions.

This study should not, however, be taken too literally as a design for an online environment. Some of these behaviors are deeply tied to the physicality of the workspace, e.g. managing placement and orientation of materials, using both hands for gesturing, or touching to get one's attention. A direct implementation of this environment would have difficulty reproducing this physicality and at the same time fail to take advantage of abilities afforded by the electronic medium (Dillenbourg & Traum, 1999; Hollan & Stornetta, 1992). Instead, we should recognize that communicative and representational practices emerge from interaction. This study demonstrates the emergence of powerful, context-specific mechanisms in underconstrained collaborative interaction. As Suthers (2005) proposes, the appropriate next step is to iterate over a series of software designs, evaluating each in terms of a qualitative understanding of how the affordances each makes available influence and facilitate users' collaborative processes.

People make sense of the world through interactions based on their perception of how objects allow them to act, not necessarily a contemplation of the objects' purpose. Action occurs in the instant-to-instant process of perceiving and responding to the world, not in an abstract vocabulary of models and operations. Providing users with complex manipulation tools might, in fact, be less important than providing effective context and information presentation (Button & Sharrock, 1997). In the long run, one of our challenges is to determine how

collaboration technologies can make the best use of computational resources and most effectively leverage the processes of sense-making already taking place in the minds and interactions of our users.

## ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under award 0093505. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

- Button, G., & Sharrock, W. (1997). *The production of order and the order of production*. Paper presented at the ECSCW '97, Lancaster.
- Conklin, J. (2003). Dialog mapping: Reflections on an industrial strength case study. In P. A. Kirschner, S. J. B. Shum & C. S. Carr (Eds.), *Visualizing argumentation: Software tools for collaborative and educational sense-making* (pp. 117-136). Great Britain: Springer-Verlag.
- Dillenbourg, P. (in press). Designing biases that augment socio-cognitive interactions. In R. Bromme, F. Hesse & H. Spada (Eds.), *Barriers and biases in computer-mediated knowledge communication—and how they may be overcome*. Dordrecht: Kluwer.
- Dillenbourg, P., & Traum, D. (1999, December). *The long road from a shared screen to a shared understanding*. Paper presented at the Proceedings of the 3rd conference on computer supported collaborative learning, Stanford.
- Garfinkel, H. (2002). *Ethnomethodology's program: Working out Durkheim's aphorism*. Lanham, MD: Rowman & Littlefield Publishers.
- Herring, S. C. (1999). *Interactional coherence in CMC*. Paper presented at the Proceedings of the 32nd Hawaii International Conference on System Sciences, Hawai'i, USA.
- Hollan, J., & Stornetta, S. (1992). *Beyond being there*. Paper presented at the Proceedings of CHI'92 Human Factors in Computing Systems.
- Kato, H., Yamazaki, K., Suzuki, H., Kuzuoka, H., Miki, H., & Yamazaki, A. (2001). Designing a video-mediated collaboration system based on a body metaphor. In T. Koschmann, R. Hall & N. Miyake (Eds.), *CSCL 2: Carrying forward the conversation* (pp. 409-423). Mahwah, NJ: Lawrence Erlbaum.
- Nilsen, E., Jong, H., Olson, J. S., Biolsi, K., Reuter, H., & Mutter, S. (1993, April 24-29). *The growth of software skill: A longitudinal look at learning & performance*. Paper presented at the InterCHI '93, Amsterdam, the Netherlands.
- Norman, D. A. (1988). *The psychology of everyday things* (1st ed.). New York: Doubleday.
- O'Neill, J., & Martin, D. (2003). *Text chat in action*. Paper presented at the GROUP'03, Sanibel Island, Florida, USA.
- Rittel, H., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, 4, 155-169.
- Sacks, H. (1984). Notes on methodology. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 21-27). Cambridge: Cambridge University Press.
- Selvin, A. M. (2003). Fostering collective intelligence: Helping groups use visualized argumentation. In P. A. Kirschner, S. J. B. Shum & C. S. Carr (Eds.), *Visualizing argumentation: Software tools for collaborative and educational sense-making* (pp. 137-163). Great Britain: Springer-Verlag.
- Shipman, F. M., III, & McCall, R. (1994). *Supporting knowledge-base evolution with incremental formalization*. Paper presented at the CHI'94, Boston, MA.
- Suthers, D. (2005). Technology affordances for intersubjective learning: A thematic agenda for CSCL, *To appear in Proceedings of Computer Supported Collaborative Learning 2005, Taipei, May 30-June 4th 2005*.
- van Bruggen, J. M., & Kirschner, P. A. (2003). Designing external representations to support solving wicked problems. In J. Andriessen, M. Baker & D. Suthers (Eds.), *Arguing to learn: Confronting cognitions in computer-supported collaborative learning environments*. Dordrecht: Kluwer.