Abstract

Real-time text and IM

By

Jacob Solomon

Advisor: Mark Newman

Real-time, character-by-character display of messages in a text-based synchronous communication system such as instant messaging (IM) can potentially lead to better communication than more traditional styles of text-display by simulating spoken dialogue through text. This research has attempted to empirically verify that real-time text display affords a truly better simulation of spoken dialogue than the standard style of text display, in which keyboard status information is displayed to interlocutors until a message has been completed. Three experiments have been conducted to look at the effect of real-time text on turn-taking, typing ability, affect towards the interface, collaborative completion, self-editing, information-sharing and communication effectiveness, and multitasking. In the first experiment, it was observed that real-time text led to better turn-taking and less self-editing. Real-time text did not influence typing ability or affect, and collaborative completion was not well supported. In the second experiment, the use of real-time text did not result in better information sharing and problem solving in a hidden profile task, although some peculiarities in the data raised new questions about how people communicate when using real-time text. In the third experiment, real-time text did not deter from effective multitasking. From this data, important design considerations for text-based CMC have been outlined.

Real-time text and IM

By

Jacob Bennion Solomon

A thesis submitted in partial fulfillment
of the requirements for the
Masters Thesis Option Program
School of Information
University of Michigan
2010

Thesis committee:

Assistant Professor Mark Newman

Research Associate Professor Stephanie Teasley

Contents

Introduction

Instant messaging (IM) and chat applications have become crucial media for inter-personal communication and for collaboration in academic and business settings. Numerous protocols and interfaces exist which enable short messages to be communicated quickly and presented in a way which supports rapid, text-based conversation online. One feature found in many early chat interfaces such as UNIX Talk was that messages were displayed in nearly real time. The recipients would see each character as it was typed. Interfaces evolved to begin only sending messages over the line after the user hit the "Return" key, and also offering some indication of the status or keyboard activity of each user. This type of interface has been the standard for IM systems for nearly ten years. Figure 1 demonstrates the difference between “traditional” IM where only completed messages are sent across, and IM incorporating real-time, character-by-character text display.

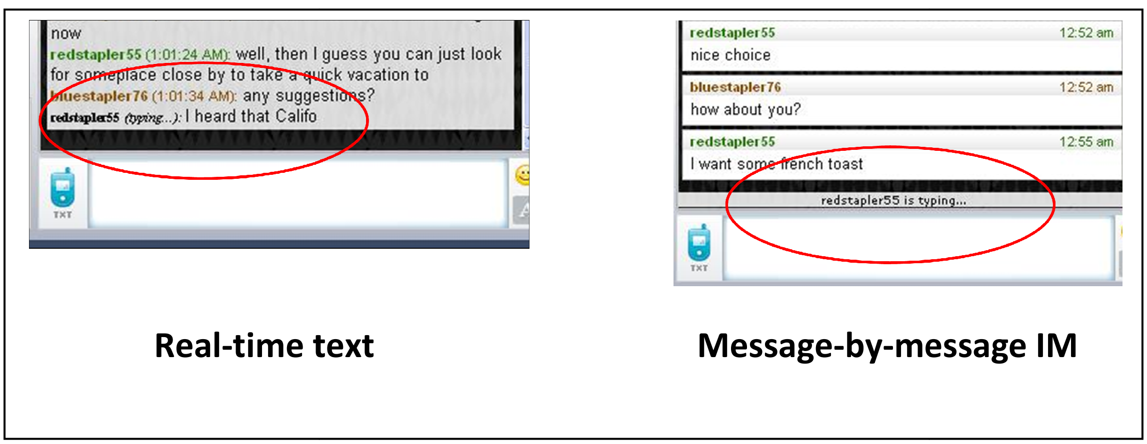

Figure 1. A comparison of real-time, character-by-character transmission of text to traditional, message-by-message transmission accompanied by keyboard status information.

Real-time display of messages has recently been reintroduced into important systems such as the America Online Instant Messenger (AIM) and Google Wave. Additionally, real-time text has been advocated as an important accessibility feature of telecommunication (“Real-Time Text Taskforce,” n.d.). It has been suggested that real-time text affords a more “conversational” style of communication than standard instant messaging ("Real-time text versus...," n.d.) and allows users to have "faster conversations" ("About Google Wave," n.d.). These assumptions have clear foundations in research on the differences between computer-mediated and face-to-face communication as well as the differences between spoken and textual discourse. A comparison of various methods of communication (H.H. Clark & Brennan, 1991) explains the characteristics of IM or chat in comparison to other communications media and to face-to-face interaction. Traditional instant messaging claims elements of cotemporality and sequentiality which are shared by face-to-face communication, as well as other mediated forms of vocal communication such as telecommunication and video conferencing. These media can be categorized separately from email and other asynchronous media which do not support a conversational style of communication. However, extensive research on instant messaging and chat has revealed that the synchronicity and sequentiality of instant messaging are not equivalent to those same characteristics in verbal communication (Berglund, 2009; Herring, 1999; Voida, Newstetter, & Mynatt, 2002; Vronay, Smith, & Drucker, 1999). This disparity has been explained by the lack of information about an utterance while it is in production, leading IM to be termed a "quasi-synchronous" medium (Garcia & Jacobs, 1999). Real-time text presents an opportunity to remove this disparity and bring instant messaging closer to verbal conversation by adding the information about message production which is lacking in traditional message-by-message IM. In doing so, real-time text can in theory afford synchronicity and sequentiality which more closely simulate verbal communication. Additionally, real-time text to some extent impairs the feature of reviewability, as users must repair their utterances publicly should they type something incorrectly or change their mind about what they want to say. The impaired reviewability makes real-time text an even closer match to telephone conversation under Clark and Brennan's framework. Thus, it is clear that there is a sound theoretical foundation for the assumption that real-time text is a better simulation of verbal conversation than standard instant messaging interfaces.

This foundation, however, lacks empirical evidence from the fields of human-computer interaction (HCI), computer-mediated communication (CMC), or computer-supported collaborative work (CSCW). Some research exists comparing real-time text to archaic forms of chat interfaces (Hancock & Dunham, 2001 ; Phillips, 2000; Vronay et al., 1999), but there is not yet a comparison of real-time text to modern IM interfaces which have made great advancements in presenting information about user availability, keyboard activity, and topic threading. It hasn't yet been empirically verified that real-time text truly offers a better simulation of spoken conversation than a modern IM client. One of the primary goals of the present research is to investigate that question.

IM has become a powerful tool for collaboration in the workplace, among researchers and academics, and as a tool for general interpersonal communication (Grinter & Palen, 2002; Herbsleb, Atkins, Boyer, Handel, & Finholt, 2002; Isaacs, Walendowski, Whittaker, Schiano, & Kamm, 2002). The usability of any type of interface in those contexts is of critical importance and can have far-reaching impact on the users' ability to function efficiently. Researchers in HCI have sought to develop heuristics, frameworks and methods which assist in the design of systems that support efficient and productive work. The usability of interfaces is typically measured by the level of productivity experienced when using the interface, such as the rate at which tasks can be completed and the quality of the work produced. Additionally, user interfaces are often evaluated as a product of user satisfaction and affect. The extent to which users enjoy using an interface is a critical consideration which can impact the adoption of technologies, as well as the effectiveness of such technologies (Dillon, 2001). Like any technology, the design of the user interface for an IM client will impact its usability. The inclusion of real-time text in an IM interface is a design decision which could affect the user experience of the interface, as well influence the type of communication which is mediated by the interface. Regardless of whether or not real-time text is more like spoken dialogue, its inclusion in an interface design may have consequences which detract from the overall usability of IM, or which may result in unpredicted usability benefits. Another goal of the present research is to evaluate real-time text as a specific design feature of an IM interface and determine the effect it has on the overall usability of that interface.

A third goal of this research is to take what is learned about the nature of communication with real-time text and the user experience of its interface and develop design implications for collaborative tools. By looking in detail at how real-time text affects communication and usability, insights can be achieved into making contextually appropriate implementations of real-time text both in IM clients and in other types of systems. These insights will be refined by the analysis of the usability and user experience of real-time text.

To summarize, this research has sought to examine the extent to which real-time text simulates spoken conversation, determine the effect of this simulation on the overall usability of IM and subjective user preference for or against real-time text, and propose design considerations for collaborative tools from the results of these analyses.

To achieve these goals, specific research questions were formulated after a review of the existing literature on chat, IM, and real-time text. As noted above, the problem of sequentiality in IM and chat has been well documented. Both ethnographic research (Berglund, 2009; Voida et al., 2002) and experimental research (Phillips, 2000; Vronay et al., 1999) have demonstrated this problem. A critical question which needs to be answered when comparing real-time text to spoken conversation is the effect it has on users' ability to coordinate turn-taking within their discourse. Additionally, it is important to measure basic characteristics of the discourse, such as the number and length of turns, in order to evaluate any differences in the style of discourse. Collaborative completion, in which interlocutors finish each others' sentences (by either explicitly saying something to finish the utterance, or simply be mentally predicting how an utterance will end), is also a feature which could potentially be supported by real-time text. If collaborative completion is supported, it offers evidence of a more speech-like style of communication.

Because IM is so prevalent in collaborative work, the effect of real-time text on the quality and efficiency of communication has important design implications. Hidden profile tasks have been used to for this purpose in previous research. (Wittenbaum, Hollingshead, & Botero, 2004) A measurement of users' ability to share information and collectively solve a problem can offer a strong argument for the usability of real-time text and for its inclusion or exclusion on the design of collaborative tools.

Another important usability consideration is the effect real-time text has on productive multitasking. Multitasking while communicating over IM has become a widespread habit among internet users, especially among teens (Lenhart, Hitlin, & Madden, 2005). A more engaging, conversational style of communication may in fact detract from IM's usability when multitasking due to increased demands on attention. Answering this question can inform the design of IM and chat clients, as well as other communicative tools which afford simultaneous tasks, such as collaborative editing tools or customer service chat sessions where a single operator may need to communicate with a customer while working separately to assist that customer, or who may need to communicate with several customers simultaneously for efficiency.

Real-time text almost certainly alters the experience of an IM user who is the recipient of an incoming message because of the additional information it provides. The experience of the message producer should also be considered in this evaluation. Real-time text not only reveals the content of a message in production, it also reveals information about *how* the message is being produced. The speed and accuracy at which it is typed are both made available to all participants, as are any repairs or revisions to the utterance. This is a natural and comfortable part of spoken dialogue. Speaking is likely a more refined skill than typing even for expert typists. So it is important to consider how people react to having their typing skills on display to their collaborators just as they put their oratory skills on display whenever speaking with someone.

Performance metrics associated with the above questions can be insightful when considering new designs for IM clients or collaborative tools. However, the usability of any interface design includes users' affect towards the system in addition to their explicit performance metrics (Dillon, 2001). The adoption of any new design will naturally be influenced by how much people *enjoy* using the interface. This approach to evaluating usability cannot be ignored for an IM client or for real-time text as a specific feature within that client. An analysis of how users feel about real-time text and why they feel that way is necessary for an optimal design.

These research questions can be summarized as follows:

1. Will the use of real-time text result in better turn-taking by IM users?
2. Will the use of real-time text help users communicate more effectively by improving information-sharing in a hidden profile task?
3. Will real-time text alter users' ability to perform another simultaneous task while using IM?
4. Will the public display of each user's typing ability affect their typing performance or their style of message production?
5. Will users enjoy using real-time text more than message-by-message IM, and what factors contribute to their emotional reaction to the interface?

To investigate these questions, three studies were designed and conducted at the University of Michigan. The first study was designed to evaluate turn-taking coordination, typing ability, and gather data about affect and the overall usability of real-time text. This study revealed that users do a better job of coordinating turns when using real-time text. It also revealed that users make fewer self-edits when using real-time text. This study found that real-time text did not influence the typing ability of participants. There was evidence that real-time text supports implicit collaborative completion, as users did try to mentally predict how utterances would end. However, there was very little explicit collaborative completion. This study found that real-time text is not significantly more preferable than message-by-message IM, however, it found a number of factors which contribute to users' affect for the system, including their experience with IM and the ratio of how much they typed to how much they read. The second study implemented a hidden profile task to learn about information sharing and overall effectiveness of communication through real-time text. This study found that real-time text did not lead to a better outcome in solving the problem. However, it also found that information sharing was less critical to that outcome when using real-time text than when using message-by-message IM. The third study looked at the effect of real-time text on multitasking, and found that real-time text did not distract users from performing a simultaneous task.

Study one: Turn-taking, typing ability, and affect

The first step in examining real-time text and its relationship to spoken conversation was the identification of the important characteristics of spoken conversation which may be facilitated by real-time text. Turn-taking coordination, collaborative completion, and inhibition to self-edit utterances were identified as characteristics of spoken conversation that could be measured in an IM conversation using real-time text. This first study found that real-time text leads to better turn-taking coordination, less self-editing, and that collaborative completion is indeed possible using real-time text, although not well facilitated. In looking at the influence of real-time text on the usability of IM, typing ability was not found to be influenced by real-time text in this study. Additionally, there was not a clear favorite among participants between real-time text and message-by-message IM, although it was found that affect towards one style or the other was influenced by previous experience with IM. A user's ratio of typed-to-read utterances also influenced which style they preferred, with heavy "talkers" preferring message-by-message and heavy "listeners" preferring real-time text.

Turn-taking

Turn-taking has been found to be a significant challenge in computer-mediated communication (Berglund, 2009; Garcia & Jacobs, 1999; Herring, 1999; Neill, 2003; Voida et al., 2002). Levinson (1983) defines ordinary conversation in terms of sequences of adjacency pairs (Question-Answer, Statement-Response, Greeting-Greeting etc.). Instant messaging systems have been found to inhibit turn-taking because one participant will submit a message which doesn't act as an appropriate second pair part to the most recently posted first pair part (Phillips, 2000). For example, if one participant asks a question and the next posted turn is a statement which doesn't respond to the question, but rather begins a new adjacency pair or responds to an earlier question, the turn-taking system has been disrupted. IM has been termed a "quasi-synchronous" medium (Garcia & Jacobs, 1999) because "message production process is available only to the person composing the message." This poses problems in CMC because the turn-taking system is based on cues made available as part of the production of messages (Sacks, Schegloff, & Jefferson, 1974). Previous research examining chat room interfaces and instant messaging has confirmed that users of quasi-synchronous CMC systems have problems taking turns (Hancock & Dunham, 2001; Phillips, 2000; O'Neill & Martin, 2003; Vronay et al., 1999).

Real-time text presents a potential solution to the problem of turn-taking by making message production visible to the message recipient. Research comparing quasi-synchronous CMC to real-time text has confirmed that real-time text can improve turn-taking (Phillips, 2000; Vronay et al., 1999). However, this research made the comparison to systems which present no information about message production to recipients. Modern IM systems offer forms of ambient notification of participants' keyboard status and availability. Most IM systems notify message recipients when the message producer is typing or has entered text. Sacks et al. (1974) point out that the turn-taking system consists of two techniques: the current speaker selects the next speaker and the speaker is self-selected. Quasi-synchronous CMC systems do not afford self-selection since not all participants are aware that a speaker is occupying the next turn while they are producing it. But modern interfaces have this affordance, and real-time text only offers to provide information about the content of the message in addition to information about the occupant of the turn. This study has sought to examine whether the real-time display of message content can lead to improvement in turn-taking in comparison to ambient keyboard status notification.

One potential problem with seeing message production in a text-based CMC system is that messages are typed much slower than they can be read. One of this study's hypotheses is that real-time text could lead to a new type of communication error in IM in which users begin predicting the final content of a message while it is under construction, and will begin formulating a response based on their predictions as opposed to the actual delivered message. Collaborative completion, or finishing an utterance initiated by someone else, is a known phenomenon in spoken conversation (Lerner, 2004) and real-time text affords this characteristic of verbal conversation in IM. Real-time text could lead to increased backtracking and clarifying of hastily submitted responses when recipients' incorrectly predict incoming messages. Real-time text could in fact make collaborative completion a frequent occurrence if users are frustrated by the relatively slow rate at which text appears on their screen.

Speech overlap, where more than one person is speaking simultaneously, has been used in conversational analysis as a measure of the functionality of a turn-taking system. (Levinson, 1983; Sacks et al., 1974). Levinson (1983) reports that in spoken conversation, less than 5% of the speech stream involves overlap. The amount of overlap in the speech stream of instant messaging conversations can similarly be measured to evaluate the turn-taking system in IM. A comparison of overlap in the real-time text condition to both message-by-message IM and phone conversations can reveal the extent to which real-time text simulates spoken conversation in comparison to message-by-message IM. Phone conversations offers a better point of comparison than the face-to-face comparison recorded by Levinson because they do not transmit as much backchannel nonverbal information as is available in face-to-face conversation. Shriberg, Stolcke, and Baron (2001) found that in phone conversations which excluded overlap due to backchannel communication (i.e. “uh-huh”), nearly 8% of the speech stream involved overlap. In a study of overlap in instant messaging, Campbell (2004) found that approximately 17% of the text stream was overlapped between participants. For real-time text to credibly demonstrate a more speech-like experience, the rate of overlap should move away from that of message-by-message IM and fall more closely to what has been found in mediated verbal conversation.

Typing

Vronay et al. (1999) observed that participants in their group chat experiment typed more accurately and quickly when using an interface which incorporated real-time text than when using a standard chatroom. They explained this finding as the result of social pressure to appear competent at typing and at communicating in general. The authors believed that participants simply tried a lot harder to type well when others could see their typing, and that in most cases this resulted in both improved speed and accuracy. They noted one participant who demonstrated an incredible increase in typing ability in the real-time text condition, but who also contributed far fewer turns to the conversation in exchange for this improvement. When interviewed, this participant cited anxiety over having his typing ability displayed for the decrease in overall participation.

The type of anxiety felt by this participant is an important usability consideration for real-time text. Voida et al. (2000) describe a “persistence and articulateness tension” in IM in which users show concern for how their spelling and articulateness is perceived. Because of the persistence of the text on the screen, users suddenly become embarrassed by their mistakes after they have sent the message across. The authors describe numerous instances of apologetic behavior in IM conversations for poor spelling or grammar. Campbell (Campbell, 2005) conducted an experiment to measure user awareness of typos and to measure the relative importance of correcting errors. He found that users went to great effort to correct errors before sending them across the line to their partner. Baron (2005) found through ethnographic analysis of college students' IM conversations that users pay close attention to spelling and language in IM.

Collectively these studies confirm that self-presentation is an important consideration for IM users. Users want to appear to be competent typists and spellers. Real-time text presents a dilemma to this consideration and presents an awkward situation to users who want to send well articulated and error-free messages across, but do not want to be seen making the necessary corrections required to appear articulate. The "persistence and articulateness tension" is altered slightly by real-time text. The result found by Vronay et al. seems a likely outcome; that users simply put more effort into typing, and they improve their performance as a way of navigating this tension. But the authors of that study also indicated that their result could not be generalized over time, and they wondered if eventually users would stop putting as much effort and focus into typing well.

Replicating the results of that study specifically in IM can give designers useful information about user behavior. If participants type faster and more accurately when using real-time text, it would offer evidence that an overall improvement in the effectiveness of communication can be attained by implementing real-time text. It is also likely to improve the user experience for message recipients by presenting a stream of text with less frequent interruptions from editing sequences. Based on the previous literature, it was hypothesized that users would perform better in terms of typing speed, typing accuracy, and in overall performance.

Affect

The variables listed thus far are all measures of user performance. They look at how real-time text affects the measurable outcomes of IM use. The usability of a system cannot be judged purely based on how efficient users accomplish tasks using a system or on how the system facilitates performance (Dillon, 2001). User feelings of enjoyment, annoyance, comfort, frustration, or anything else can offer important insight which may not be captured by looking strictly at performance or outcomes. In order to properly evaluate the usability of real-time text as a distinct interface feature, some data was needed which gauged how users felt about real-time text. And perhaps the single most useful piece of data in this type of usability evaluation is simply to know which type of interface users prefer. These data can capture those factors which were simply immeasurable or which were not considered. Even if they cannot explicitly parse out all the variables which influence their performance, users know whether an interface is working for them. The data about affect and preference then serve as an important catch-all in the overall evaluation of real-time text. If anything about the systems detracts from their overall usability, this will impact the way users feel about the system.

Additionally, for the purposes of informing design, it is necessary to know what users actually prefer and why. In many contexts, particularly commercial applications in which users have the choice about whether to use a system or not, designers cannot include any interface elements such as real-time text if it will cause users to reject the entire system, even if they can demonstrate greater efficiency or that the feature will lead to better outcomes. If people won't use it, then the system will have failed.

Besides knowing what users prefer and how they feel about real-time text, it is important to determine the factors which influence these attitudes and feelings. This provides an opportunity to merge the outcome-based and descriptive data collected in these experiments with the data about affect to investigate *why* users preferred one over the other and which factors contributed to their emotional response to the systems. These analyses will explain how real-time text can be implemented to maximize user satisfaction, and can also inspire modifications to the design of real-time text or even new designs altogether which can improve the usability of CMC.

Hypotheses

1. Real-time text will lead to fewer disrupted adjacency pairs.
2. Real-time text will lead to less simultaneous typing.
3. Users will report instances of collaborative completion errors.
4. Users in the real-time text condition will have more deletion sequences to make up for their collaborative completion errors.
5. Users will show improved typing ability, in both speed and accuracy.

Methodology

AIM 7.0 was used to test real-time text with 24 participants (14 male and 10 female) recruited from the campus of the University of Michigan through flyers advertising a study on instant messaging. Respondents to the flyer were asked whether they were a native English speaker, and whether they had used IM within the last year. Only native English speakers were admitted to the study so that measures of typing ability and turn-taking would not be influenced by factors related to communication through a second language. Only respondents who had used IM at least once in the previous year were admitted, assuring that the data would not be skewed by any unfamiliarity with IM as a mode of communication. Two tasks were developed to solicit ordinary conversations between randomly paired dyads, and the participants were assigned to begin the experiment with one of the four combinations of task and text display (real-time versus message-by-message IM). One of the tasks asked the dyads to develop a list of five movies they would bring with them on a weekend retreat to a cabin as well as other food and supplies they would bring on a $200 budget for ten people. This task was nearly identical to the task used by Vronay et al. (1999), with the only change being the addition of the budgeting scenario. This part was added to assure that participants had enough to talk about for at least ten minutes, since the task used in the original study was performed by larger groups. The other task asked the participants to identify three important transportation issues on the campus of the University of Michigan, and then develop a plan for making a presentation to administration on behalf of students. Appendices A and B provides the details of each of these tasks. Participants would converse for up to fifteen minutes, after which they would be given another task and the type of text display would be switched. The AIM client allowed for text display to be switched without any other changes to the interface. Participants were randomly assigned to a computer station with an IM dialog already open with another participant. Experimental sessions consisted of groups of either two or three dyads. Participants were not told which other participant in the session was their partner.

After both tasks were completed, participants were given a questionnaire which asked whether they found themselves trying to predict how the other person's utterances would end when using real-time text. They were also asked to describe any instances of problems in their communication due to incorrect predictions about the incoming messages. The responses to this questionnaire were analyzed qualitatively to develop insight into participants' experience with real-time text. This questionnaire was designed to collect qualitative data about the existence of collaborative completion and the errors which could occur as a result of mistaken collaborative completions.

Chat logs, screen recordings, and keystroke logs were used to collect data from each session. The chat logs were analyzed to evaluate the effect of real-time text on turn-taking. Descriptive statistics regarding the length of turns, number of turns, and frequency of turn changing were measured. A "turn" in this sense was defined as all the text sent to the partner by one stroke of the "Enter" or "Return" key. Turn-changing frequency was then defined as the number of turns in a conversation divided by the number of times the conversation changed speakers. The level of contribution to a conversation was defined and measured as the number turns submitted by a participant minus the number of turns submitted by their partner. Each conversation was then coded and broken into adjacency pairs. Clark's (1989) description of adjacency pair types was followed in coding the conversations. A turn was coded as a first pair part or a second pair part, and pairs that belong together were identified. When an adjacency pair was separated by a part of a different adjacency pair, it was also coded as a disrupted pair. Figure 2 demonstrates an example of a disrupted adjacency pair. The disrupting adjacency pair could come from either speaker. A disrupted pair could result from a speaker asking a question, then asking another question before their partner submitted a response to the first question. It could also occur if participants merely tried to start new adjacency pairs nearly simultaneously.

U01(5:38:23 PM):	do you like action movies?	pair 1 part 1
U02(5:38:24 PM):	so sleepy hollow and amityville horor	pair 2 part 1
U02 (5:38:26 PM):	yeah	pair 1 part 2

Figure 2. A disrupted adjacency pair.

An exception to this coding was made whenever the impeding part represented an insertion sequence as described by Levinson (1983), in which a new but logical sequence is inserted in between the two parts of an adjacency pair, such as responding to a question with a clarifying question. Figure 3 demonstrates an insertion sequence which was not coded as a disrupted adjacency pair.

U01 (5:37:16 PM):	what's your favorite?	pair 1 part 1
U02 (5:37:25 PM):	from scary movies?	pair 2 part 1
U01 (5:37:27 PM):	yeah	pair 2 part 2
U02 (5:37:40 PM):	i like sleepy hollow	pair 1 part 2

Figure 3. An insertion sequence.

The keystroke logs were used to analyze the amount of self-editing in the conversations for comparison between the types of text display. Self-editing was measured by counting the number of instances of the "Backspace" key in the keystroke logs. A sequence of one or more consecutive "Backspace" keystrokes was recorded as one instance of self-editing. Screen recordings were consulted to explain irregularities in the chat logs, as well as to obtain qualitative data about the live process as has been suggested in previous work (Garcia & Jacobs, 1999). Some dyads finished both tasks in approximately eleven minutes, whereas others did not finish both tasks in the allotted time. Because of this disparity, only the first ten minutes of each conversation was used for the quantitative analyses of turn-taking, which assured that the sample from each pair represented active participation in conversation. Paired t-tests were used to evaluate within-subjects variables which resulted from the repeated measures design.

To measure the amount of typing overlap between participants in a conversation, a script was written in Java to mark the starting and stopping points for each turn by each member of a dyad. Two types of stopping points were defined. The first type was a stroke of the “Enter” key. The second was the final stroke of a complete deletion of entered text. This was used to prevent over-measurement of overlap in instances where both participants began typing at approximately the same time, but one participant shortly conceded the turn space. This assured that the time spent waiting for the partner to finish their turn was not counted as overlap. A starting point was then defined as either the first keystroke in the sample, or the first keystroke after a stopping point. The starting and stopping points were compared between each partner and the overlap in seconds was tallied. As in the analysis of adjacency pair disruption, only ten minutes of each conversation was analyzed to assure the sample represented active engagement in the task. A paired t-test was used to compare the mean number seconds of overlap between the real-time text and message-by-message conditions.

Keystroke logs from each experiment were recorded. These logs provided a list of timestamps, measured to the one-hundredth of a second, for each keystroke performed by participants. Three metrics were used to evaluate typing ability in the experiments. Keystrokes per second (KSPS) was used as a measurement of pure typing speed. As suggested by (Soukoreff & MacKenzie, 2003), keystrokes per character (KSPC) was used as a measurement of typing accuracy. This measurement was taken by comparing the number of characters in a completed sequence to the number of keystrokes used to create that message. The definition of a sequence here is identical to what was used for evaluating overlap in the analysis of turn-taking. A sequence ended when either a participant hit enter to send the message across the line, or at the moment they deleted all of a message without sending it over. The first keystroke following the end of a sequence defined the beginning of the next sequence. The modifier keys, such as "Shift", were ignored in these analyses. This was done primarily to avoid biasing the results of participants who spent a long time holding down the "Shift" key while thinking about what to write, and to avoid counting capitalized letters as two keystrokes. Characters per second (CPS) was used as a measure of total performance. This was taken as the number of characters in a finished sequence divided by the total time taken to construct the sequence.

Paired t-tests were used to compare the typing ability of participants between the real-time text and message-by-message IM conditions.

The IBM Computer Usability Satisfaction Questionnaire (CSUQ) (Lewis, 1995) was given to participants after completing one task in the turn-taking experiment. See Appendix C for a complete text of the CSUQ. The CSUQ asks a total of 21 questions. 19 of these questions seek specific information about user satisfaction regarding overall satisfaction, enjoyment, efficiency, comfort, and productivity. Participants responded to these questions using a Likert scale, and a field for free-response comments was available for each question. The last two questions asked participants to list their favorite and least favorite features of the system. After using both real-time text and message-by-message versions of AIM, participants were also asked to state their preference for one or the other, and to provide feedback regarding their choice. Participants were also asked how much time they spend per week using IM, and which IM clients they use.

To compare responses to the Likert scale questions, a Mann-Whitney U-test was performed comparing the real-time text group to the message-by-message group. Logistical regression was used to find factors which influenced preference for one type of text display over the other. Performance metrics from the earlier analyses were also compared to look for their influence on preference. These included the typing speed and accuracy of both a participant and their partner, the disparity in the amount of contribution to a conversation (measured by the difference in number of turns between partners in a dyad), and the length of turns in both words and characters per turn. The amount of time spent by participants using IM on a weekly basis was also analyzed to determine its influence on preference.

Results

Dyads had a significantly smaller percentage of disrupted adjacency pairs when using real-time text than when using message-by-message IM, $t(11) = 4.11$, $p < .01$. Using real-time text, a median of 29% of a dyads' adjacency pairs were disrupted, compared to 39% when using message-by-message display. Table 1 shows the distribution of these data. The type of message display did not affect the number of words or characters per turn, the number of turns per ten minutes of communication, or the frequency with which dyads changed turns. There were no significant differences in any of these measures when the data was grouped by task instead of by style of text display, assuring no task related effects.

When using real-time text, users had a lower level of typing overlap than when using message-by-message IM, $t(9) = -2.83$; $p < .05$. It should be noted that two of the observations were not included in this analysis due to data corruption of the keystroke files. In the real-time text condition, an average of 10.9% of the text stream was overlapped, in comparison to 17.1% in the message-by-message IM condition.

Using real-time text reduced the amount of self-editing by participants, $t(21) = 2.32$, $p < .05$. Participants made an average of 11.8 more self-edits when using the message-by-message display than when using real-time text.

	Real-time text	Message-by-message	Paired t-test
% of disrupted adjacency pairs	29% (11.3%)	39% (9.8%)	t(11)=-4.11; p<.01
Number of deletion sequences	38.4 (13.9)	51.8 (22.83)	t(11)=3.39; p<.01
% of time spent typing simultaneously	10.8% (8.5%)	17% (8.6%)	t(9)=-2.83; p<.05
Number of turns	60.53 (19.38)	63.83 (23.36)	Not Significant
Number words	371.75 (68.21)	389.08 (68.21)	Not Significant
Average length of turns (words per turn)	7.15 (2.83)	6.90 (2.36)	Not Significant

Table 1. Study 1 results. Standard deviations noted in parentheses.

The responses to the questionnaire presented a consistent message about collaborative completion and the occurrence of prediction errors in the experiment. Users consistently reported that as messages were created on the screen before them, they would try and predict how the message would end. However, there were only a few reported instances of errors as a result of an incorrect prediction of the incoming message. Participants stated that despite their tendency to predict the messages, they avoided responding until the other person had finished their turn because they felt a responsibility to be polite or maintain etiquette.

	Real-time text	Message-by-message	T-test
Keystrokes per second	2.884	2.267	Not significant
Keystrokes per character	1.363	1.359	Not significant
Characters per second	2.119	1.730	Not significant

Table 2. Typing ability of participants.

No differences in any of the measures of typing ability were observed, as noted in table 2.

	Real-time text	Message-by-message
Preference	10 participants (42%)	14 participants (58%)

Table 3. Preference

Table 3 indicates the stated preference for each style of text display. This difference was not statistically significant according to the binomial distribution.

Results from the CSUQ are presented below in table 4. The scores represented are the median for each question and the scale ranging from 1 to 7, with 1 being "Strongly disagree" and 7 being "Strongly agree."

Question	Real-time text	Message-by-message
Overall, I am satisfied with how easy it is to use this system	6	6
It was simple to use this system	7	7
I can effectively complete my work using this system	6	6
I am able to efficiently complete my work using this system	6	6
I feel comfortable using this system	7	7
It was easy to learn to use this system	7	7
I believe I became productive quickly using this system	5.5	6
Whenever I make a mistake using the system, I recover easily and quickly	5.5	6
The interface of this system is pleasant	5	6
I like using the interface of this system	6	5
Overall, I am satisfied with this system	6	6

Table 4. CSUQ median responses.

The Mann-Whitney U-tests found no differences on any single question from the CSUQ.

Logistic regression found that the amount of time participants used IM in a week was a strong predictor of their preference for style of text display, ($p < .05$). The more participants used IM, the more likely they were to prefer message-by-message IM.

Logistic regression also found that the difference in the number of turns “spoken” in a conversation between two partners predicted preference, ($p < .05$). The more turns a participant contributed relative to what their partner contributed, the more likely that participant was to prefer message-by-message IM.

Factors which were not found to have significant influence on preference were the length of turns, the number of turns, the typing speed and accuracy of the user or of their partner, or gender.

Qualitatively, real-time text was frequently cited as one of “the most negative features” of the system in the free-response section of the CSUQ. Some important examples were:

- “I didn't like seeing what was being typed before it was sent to me. When people make errors and have to go back and fix it it becomes very annoying to watch. I felt that it required more attention to what the other person was writing while they were writing it, instead of just reading their final messages.”
- “On AIM, it is annoying to see what someone is typing before they're done typing. I would rather see complete ideas.”
- “I did not like how it showed what you were typing as you typed it, it makes it too easy for the other person to respond to you before you've finished typing and messes up the flow of the conversation”

There were a few instances where real-time text was cited as one of the most positive features.

Discussion

These data offer evidence that real-time text helps users coordinate turn-taking. The increased information about message production, namely the content of the message itself, is more helpful to users than information about keyboard activity alone. By knowing in detail what message is incoming, users can plan their turns better both semantically (as shown by the lower percentage of disrupted adjacency pairs) and temporally (as shown by the decreased amount of simultaneous typing).

It is noteworthy, however, that despite the improvement, participants still disrupted almost 30% of their adjacency pairs. The turn-taking system in spoken conversation is dependent on interlocutors acknowledging the type of turn which has preceded (Sacks et al., 1974), and therefore this type of error is rare in spoken conversation. There are two possible and non-exclusive explanations for the high rate of disrupted adjacency pairs in the real-time text condition. One explanation could be that having information about message production is insufficient for coordinating turns in a distributed CMC context, and that other backchannel or non-verbal information such as intonation, visual cues, and timing mechanisms are required. Another explanation could be that precise turn-taking is simply not highly important to the goals of the interlocutors and that the convenience of being able to be both message producers and receivers simultaneously is more useful, and thus, users ignored turn-taking coordination when they felt it was unnecessary. A detailed qualitative analysis of these data could generate new insight into the coordination of turn-taking with real-time text. Additionally, the rate of simultaneous typing, 10.8%, is considerably higher than the 5% rate which has been observed in spoken dialogue (Levinson, 1983), although it does come closer to matching the 8% which has been recorded in spoken conversation over the telephone (Shriberg et al., 2001). This could be explained simply by the inherent difference in the difficulty of typing versus speaking. Because typing is slower and requires more effort and possibly more overall attention, there might be an incentive for a user to begin typing their message as soon as possible in order to facilitate efficient conversation. At the present moment, it can be summarized that real-time text leads to a slight improvement in turn-taking coordination. However, the turn-taking problem in IM is not completely resolved by real-time text. Therefore, further discussion on the topic should focus on weighing the slight improvement in turn-taking coordination against other influential factors. Among these factors is the overall importance of turn-taking to a CMC system, as well as other usability factors which have investigated in this research's other two experiments. Detailed discussion of these topics will be presented in the conclusion of this document.

Users in this experiment were not influenced by real-time text to type with better speed or accuracy. The within-subjects design of this experiment strengthens this conclusion, as it shows that users did not alter their own typing ability at all when switching between styles of text display. One important consideration, however, is that measuring typing speed and accuracy has been found to be exceptionally difficult in situations where users are composing their messages as they type (Soukoreff & MacKenzie, 2003). Studies evaluating typing or text entry ability are typically done using a transcription task instead of tasks where users must produce ideas or content as they go. It is very difficult to control for any variations in typing speed and accuracy which are the result of the user simply thinking about what to write or going back to change something they've written simply because they have changed their mind. Despite the evidence that users want to present articulate messages with no spelling or grammar mistakes, it is also difficult to define a completely valid measure of typing accuracy for IM. Emerging behaviors in online communication such as the use of abbreviations and acronyms (Baron, 2004) can make it difficult to determine what constitutes a "mistake" in instant messaging. For this reason, this study only used metrics which account for corrected errors. Uncorrected errors would not show up in these analyses. This is an important consideration because real-time text may in fact influence the balance between corrected and uncorrected errors. Some users may decide not to take the time to correct errors when using real-time text simply because they know their partner has already seen the error. A valuable direction for future research would be to develop a framework for differentiating corrected from uncorrected errors in IM conversations, and applying that framework to these data in order to investigate any influence of real-time text on user behavior.

The most significant finding from the affect data was that there was a difference between users who did more typing and those who did more reading. Participants who did more typing liked using message-by-message IM better, while those who did more reading liked real-time text. This can likely be explained by the inherent differences in the two types of text display. Message-by-message IM affords more freedom to edit, as discussed in the section on typing ability. Therefore, it is understandable that participants who compose a lot of messages would prefer to have the flexibility to be articulate and send well-written, error-free messages across the line. Conversely, it is understandable that users who do more reading would prefer real-time text, since it is likely more interesting to observe even parts of a message than a generic message such as "*user is typing…".*

Since the CSUQ failed to distinguish real-time text from message-by-message IM in any of its categories, it can be concluded that real-time text is not viewed by users as a factor which entirely determines the usability of an IM system on its own. It is clear from these data that the usability of an IM client depends on a variety of factors, and real-time text alone is not a strong enough feature to have a significant impact on how users feel about an interface. It should also be noted that the responses were generally quite high and favorable to the AIM client across all questions. This may indicate a problem with the instrument. The CSUQ is a questionnaire which is designed to be applicable to all types of computer systems, however, it may not be precise enough to distinguish very simple interfaces such as an IM client. Because the tasks required no interaction with the interface outside of typing and reading messages (and did not require users to change settings, connect to other participants, detect presence or availability etc.) it may simply have been that the CSUQ was too blunt and general to detect differences at such a low level of interaction.

What can be said is that users react differently to real-time text depending on the context in which they are using it, and so a general statement that real-time text has either better or worse usability overall is inappropriate. The way that users feel about real-time text is dependent on variables such as their experience and familiarity with IM, the reasons they are using IM and the way in which they are using. This is an important conclusion which confirms that an evaluation about including real-time text in a system design must take contextual factors into account.

Study two: Effectiveness of communication

As the use of instant messaging grows in the workplace (Lenhart & Shiu, 2004), it is increasingly vital that its role in facilitating effective and efficient collaboration be critically evaluated. The design and implementation of information and communication systems can be informed by exploring the nature of collaborative work over IM. This study has looked at the effect of real-time text on the effectiveness of communication through IM, and found that real-time text did not lead to better communication, although the data did unexpectedly suggest some new questions about the qualitative effect of real-time text on communication.

Ethnographic studies of IM in the workplace have reported that IM is most often used for scheduling tasks, quick conversations or for informal social interactions (Isaacs et al., 2002). IM has become a powerful tool for these kinds of interactions because it is efficient, quick, and less interruptive than other forms of communication (Nardi, Whittaker, & Bradner, 2000). Nardi et al. (2000) observed that one of IM's most useful functions was to coordinate face-to-face meetings or phone conversations. This would suggest that IM has tremendous value in an organization for presenting information about the availability of its members and for facilitating very simple interactions, but that other media (or no media) are preferred for more complex problem-solving contexts or for in-depth collaboration. This apparent lack of adoption for formal collaboration might be explained by research comparing computer-mediated communication to face-to-face communication for solving hidden profile tasks. Hidden-profile tasks are group problem-solving scenarios where each member is given only a portion of the necessary information required to solve the problem, but collectively the group holds all the necessary information. Solving the problem becomes a function of efficiently sharing the appropriate information. The superiority of face-to-face communication in these tasks (Wittenbaum et al., 2004) is likely felt intuitively by collaborators in the workplace, who then opt for richer forms of communication with more backchannels for information exchange.

Real-time text may provide an opportunity for instant messaging to become a more powerful tool for complex and formal interaction and collaboration. Because of the observed improvement in turn-taking and the potential for a closer simulation of spoken conversation, it can be hypothesized that using a real-time text IM client can lead to better problem solving and more efficient information sharing. Because users will have fewer problems coordinating exchanges and because the information exchange is more immediate, users should be able to share information quickly and efficiently. This should then lead to better problem solving in a hidden profile task.

Hypotheses

1. Using real-time text will lead to a solving the problem more accurately than message-by-message IM.
2. Using real-time text will lead to making a faster decision than message-by-message IM.
3. Using real-time text will lead to more discourse (measured by the number of characters, words, and turn)
4. The amount of discourse in a conversation will predict solution quality for both types of text display.

Methodology

To evaluate the quality of communication afforded by real-time text display, a hidden-profile task was created which required dyads to work together to solve a problem. 24 participants were recruited from the campus of the University of Michigan to participate in the experiment. Participants were screened for previous IM use and native language. Only native English speakers who had used IM at least once in the past year were allowed to participate. The participants were randomly assigned to either the real-time text condition or the message-by-message condition. Participants were randomly assigned to a computer station with an IM window open and connected to another participant. Participants were not told to which computer station they were connected, meaning participants did not know with whom they would be communicating. The participants were given instructions on paper with information about the task, and given five minutes to read the information before beginning the task with their partner. The hidden profile task used in this study was a murder-mystery task designed for dyads, similar to the task used by Meier and Spader (2007) in a study of information sharing. See appendices D and E for detailed task sheets. Participants were assigned to either the role of an FBI agent or a Police Detective and told they were to collaborate to solve the case of Mr. Body's murder based on the evidence each had collected separately. Participants were to make an arrest in the case based on three criteria; access to the murder weapon, lack of an alibi, and a motive. There were a total of forty pieces of relevant information about the case. There were twelve pieces of shared information, and each participant held fourteen pieces of unique information. Participants were told to collaborate and to make an arrest as soon as they felt confident. After deciding who to arrest, participants were asked to rank the remaining suspects according to the likelihood they were guilty. If participants shared all available information, they should have been able to rank the suspects as follows:

1. Colonel Mustard- Colonel Mustard had access to the murder weapon, his alibi was false, and he had a strong motive.
2. Professor Plum was the next best choice. He had access to the murder weapon and his alibi was false. There was only weak evidence of a motive however, that Mr. Body was Professor Plum's accountant. Each of the other three suspects also employed Mr. Body as an accountant, but had additional motives to potentially murder him (involvement in a money laundering scheme, evidence of tax evasion, and suspicion that Mr. Body was stealing money).
3. Miss Scarlet was the third best choice. She had access to the murder weapon and a strong motive, however there was also evidence verifying her alibi.
4. Mrs. Peacock was the worst choice. She had a strong motive, but there was no evidence she had access to the murder weapon, and there was evidence verifying her alibi.

The time taken to reach a decision was recorded. Participants were allotted a maximum of fifteen minutes to reach a decision. Each participant recorded their own decision. There was no mandate that they come to an agreement. Each of their rankings was awarded a score based on the proximity of their answer to the correct choice for that level (guilty, 2^{nd} most likely, 3^{rd} most likely, least likely). For example, if a participant chose Professor Plum as the guilty party, they would receive one point for that choice because Professor Plum was the second best choice, meaning he sits at a distance of one from the chosen location. If this user also chose Miss Scarlet as the second best choice, they would receive another point because Miss Scarlet should be the third best choice. Choosing Colonel Mustard at third would add two points to the score etc. Their total score would then be the sum of all their scores. A perfect score would be 0. This approach of collecting a decision about each suspect was used to assure that the results more closely reflected the total amount of information shared instead of a measurement of who happened to discuss the information relevant to the actual murderer. In other words, accurately establishing conclusions about each criterion for each suspect was treated equally. T-tests were used to compare the mean scores and times between the real-time text group and the message-by-message group. An analysis of covariance was also performed comparing each group's performance relative to the amount of time spent working on the problem. This analysis was intended to account for any differences in problem-solving performance resulting from a difference in the amount of time spent working on the task.

Using outcome-based evaluations of group tasks in CMC as a solitary measure of group performance has been criticized due to a high number of confounding variables which can be difficult to control for (McCarthy & Monk, 1994). For this purpose, the number of words, characters, and turns used in each conversation were recorded and used as another measurement of information sharing. These metrics were also compared between groups using t-tests. To evaluate the validity of these metrics as being representative of true information sharing, linear models were fit to determine the correlation between information-sharing and problem-solving.

Results

	Real-time text	**Message-by-message**	**T-Test**
Score	3.58 (2.31)	3.33 (2.74)	Not significant
Time (seconds)	731.8 (174.14)	806.2 (151.72)	Not significant
Characters	2561 (539.58)	2700 (1157.69)	Not significant
Words	499.6 (102.26)	561.8 (157.96)	Not significant
Turns	60.4 (11.08)	59.7 (15.8)	Not significant

Table 5. Study 2 results. Standard deviations in parentheses.

The t-test found no difference between the real-time text group and the message-by-message group in terms of their scores on solving the problem. There was also no difference between the conditions in the amount of time taken to solve the problem. The analysis of covariance found no interaction effect of time and performance. Table 5 summarizes these results.

There was no difference between the two groups in the number of turns, characters, or words used to solve the task.

Within the message-by-message group, the number of turns used was highly predictive of problem-solving performance, $F(1,10) = 11.81$; $p < .01$, meaning dyads who sent more messages over the line were better at solving the problem. This finding is replicated when measuring information sharing in terms of the number of characters sent over the line, $F(1,10) = 7.473$; $p < .05$. When measuring information at the word level, a similar trend was observed, $F(1,10) = 4.287$; $p < .1$. Table 6 shows the relationships between information sharing and problem-solving.

	Real-time text	**Message-by-message**
Characters	Not significant r = .02	F(1,10) = 7.47; p < .05 r = -.65
Words	Not significant r = -.09	F(1,10) = 4.23; p < .1 r = -.55
Turns	Not significant r = -.29	F(1,10) == 11.81; p < .01 r = -.74

Table 6. Linear models using score as the response variable. Note that a high score represents poor performance in solving the problem.

Neither the number of turns, words, or characters had any influence on participants' ability to solve the problem when using real-time text.

Discussion

These results suggest that using real-time text when instant messaging is unlikely to result in noticeable improvements in collaboration, information sharing or problem solving when compared to more traditional instant messaging clients. Thus, the inclusion of real-time text in the design of an IM client cannot be justified entirely on the basis of improved communication in problem-solving or information sharing contexts. This study does not look at real-time text in other forms of CSCW or CMC such as collaborative document editing tools or virtual meeting systems, however, these results may be applicable in those contexts as well. Since those types of tools are designed specifically to aid collaboration for complex problems and contexts, implementing real-time text in place of a message-by-message style of interaction is unlikely to provide any great boost to the effectiveness of the system as a tool for collaboration. It will not serve to automatically make such a system more productive or to better facilitate remote collaboration.

The finding that more information sharing in the real-time text group did not influence problem-solving ability is somewhat perplexing, given that the same measures did influence problem-solving in the message-by-message group, and especially because there was no difference in problem-solving ability between the groups. What might be inferred is that when using real-time text, participants took a different strategy to solving the problem that was not dependent on merely exchanging as much information as possible. What is even more noteworthy is that this strategy appears to have been just as effective as the strategy taken by the message-by-message group, since the real-time text group did just as well at solving the problem. These data propose an important direction for future research regarding the influence of real-time text on the style of collaboration. Further investigation which characterizes this difference can be valuable for understanding specific contexts of collaborative work which may benefit most from an implementation of real-time text.

Study three: Multitasking

One important characteristic of instant messaging is that its use does not exclude other simultaneous activity. This study has looked specifically at the influence of real-time text on users' ability to actively engage in a separate, attention-intensive task while using IM, and found that it was not more difficult to multitask when the IM conversation used real-time text than when it used a message-by-message interface.

Multitasking is a widespread behavior among IM users (Lenhart et al., 2005). This makes IM an advantageous form of communication in many contexts, including the workplace (Nardi et al., 2000), among teens (Grinter & Palen, 2002), and in educational settings (Nasah, 2008). These studies broadly explain that instant messaging is a less interruptive form of communication than others and therefore, users have demonstrated success at performing tasks while engaging in an IM conversation without disruption to their performance on these tasks. The most apparent explanation for this is that IM affords a semi-synchronous style of communication, where attention can strategically be allocated away from the conversation when a user is not actively engaged in either composing or reading a message. The lack of backchannel or non-verbal information is beneficial in this context because cognitive resources can instead be allocated to another task.

Real-time text, by attempting to more closely simulate spoken dialogue, poses a threat to this advantageous characteristic of IM. If users are compelled to allocate more attention to an IM conversation because of its synchronous nature, they may have a harder time performing another task simultaneously. Concern over the loss of multitasking support in IM has already led to new design proposals, such as a system for controlling the saliency of an incoming message according to its importance (Avrahami & Hudson, 2004). The purpose of this experiment is to evaluate the saliency of real-time text as a distraction.

Previous experiments looking at text-based CMC and multitasking (Dabbish & Kraut, 2004; Scupelli, Fussell, Kiesler, Quinones, & Kusbit, 2007) have focused on users' ability to coordinate their interruptions and manage timing and awareness between tasks and collaborators. The intent of this experiment is to measure the demand on attention caused by real-time text and to observe any resulting effect on users' multitasking capabilities. It was hypothesized that real-time text will make multitasking more difficult than for two reasons. First of all, the visible scrolling text will naturally attract attention. This hypothesis is based on well established theories in experimental psychology demonstrating that abrupt motion attracts attention (Hillstrom, 1994). Secondly, the closer simulation to the communication conventions of spoken conversation will lead to more attention being allocated to the conversation when the IM interface uses real-time text. By demanding more attention, real-time text will therefore lead to poorer performance on a simultaneous task than message-by-message IM.

Hypothesis

1. Message-by-message IM users would perform better on the pattern recognition game and in total performance.

Methodology

28 English-speakers were recruited from the campus of the University of Michigan to participate in the experiment. Participants must have used IM at least once in the past year in order to participate. They were randomly assigned to either the real-time text condition or the message-by-message condition. Participants were not told who their partner was among the three to five other participants in the experimental session. Prior to beginning this experiment, participants performed the hidden profile experiment from study two. Participants were asked to perform two simultaneous tasks. One task was to play the computer game Anticipation. The game presents users with a window full of dots which appear to be randomly scattered across the game's screen. A countdown begins and lines are slowly drawn between the dots to form a picture. At the top of the window, a category related to the picture is displayed. For example, if a hammer is being drawn, it might say "tool" above the drawing. The object of the game is to guess what picture is being drawn as quickly as possible. As soon as the user knows the answer, they press a key to stop the countdown and enter their answer. They have a limited amount of time with which to answer after stopping the countdown. If they are correct, they receive points based on how quickly they answered. Optimal success in the game requires constant attention but otherwise very little skill, meaning that performance in the game is a strong indicator of the level of attention given to it by the user. In order to prevent any learning curve effects, participants were given instructions and allotted five minutes to play independently and become familiar with the controls and objectives of the game. Figure 5 represents a typical scenario from the Anticipation game.

Figure 5. The Anticipation game

For the second task, the party game Catch Phrase was modified to be played over IM. In this game, participants were given a list of common words and short phrases which were taken from a commercially available version of the game. Each partner had a different set of words. The objective of the game is to have one's partner guess each word from the list. To do this, a participant may say anything about the word or phrase they want without explicitly using the word or any part of it. Participants would check each word off the list as soon as their partner correctly guessed it. Participants could go through the list in any order they preferred, and dyads were free to coordinate their turn-taking however they wished. Each participant's score on this task was the sum of the words they correctly guessed and the words their partner guessed.

Participants were instructed to work for a balanced performance between the two games. They were also told that the top four total scores (the product of the Anticipation score and Catch Phrase score) would be given an additional $10. This was done to assure active and balanced participation across the tasks.

The AIM client was placed at the center of the screen, with the list of words for catch phrase to the right and the Anticipation window to the left. 27” monitors were used at a resolution of 1920x1200, which enabled all windows to be displayed at the same time without interference. The experimental session lasted 15 minutes. Figure 6 shows the screen arrangement for the task.

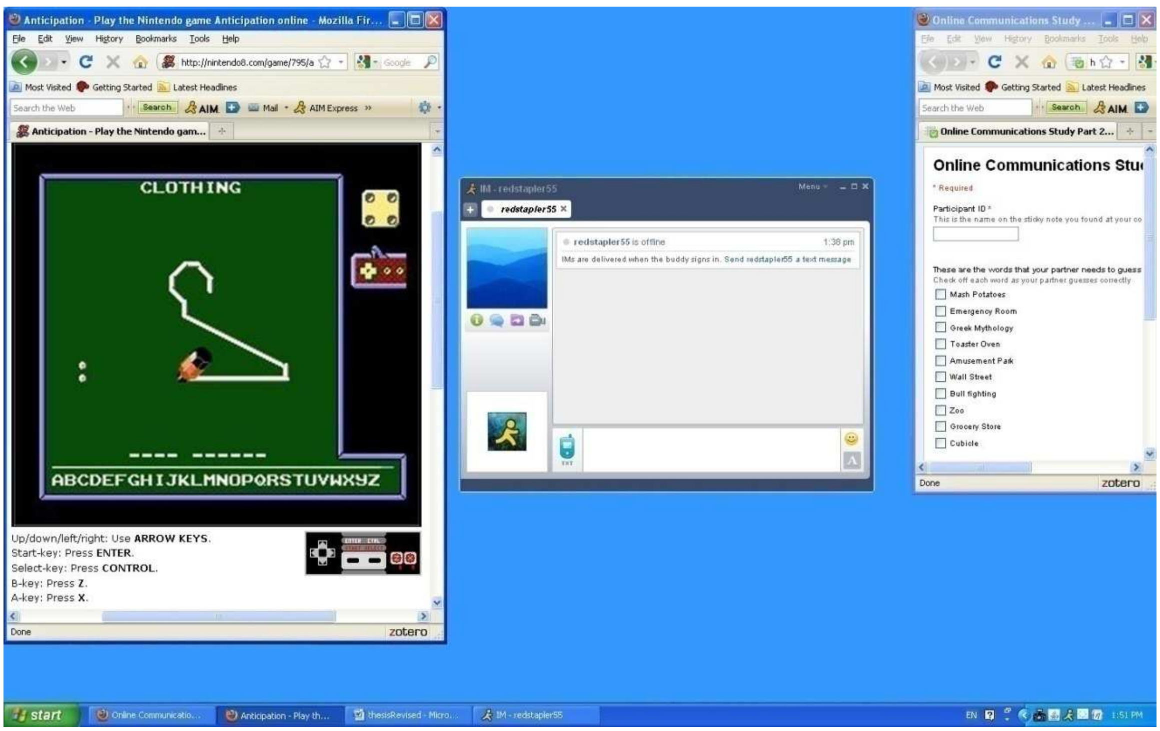

Figure 6. Screen arrangement for the multitasking experiment.

A t-test was used to compare the performance of the real-time text group to the message-by-message group on each task. An analysis of covariance was used to evaluate total performance.

Results

The t-tests found no difference between the real-time text group and the message-by-message group in their scores on the Anticipation game. No differences were found between the scores of the Catch Phrase game. The ANCOVA found no interaction effects. Table 7 summarizes the results.

	Real-time text	**Message-by-message**
Anticipation score	224 (85.48)	205 (77.51)
Catch Phrase	16.1 phrases (3.67)	15 phrases (2.70)

Table 7. Scores from the multitasking experiment. Standard deviations in parentheses.

Discussion

The experiment did not find that real-time text made it harder for users to perform another task simultaneously. The scrolling text which appeared on the IM window did not significantly distract users trying to play Anticipation. There are a number of possible explanations for this outcome. First of all, users may have chosen to ignore incoming text until it appeared as a completed message, which would have rendered real-time text to be nearly identical to message-by-message IM. Since all these users had spent fifteen minutes performing the hidden-profile task just prior to this experiment, the participants may have had the opportunity to learn that watching the text scroll onto the window at a slow rate (an average of about 2 characters per second) is an inefficient use of their time and that the best way to achieve maximum overall performance would be to play Anticipation until a message has been completed. This would suggest that if the user has an incentive to ignore real-time text, they are capable of doing that without the need to explicitly turn it off in the interface. The attraction of such small bits of new information as the first few words of a message just may not be strong enough to influence a user who is actively engaged in another activity.

There are experimental factors which may have contributed to this result. It is possible that the windows containing the Anticipation game and the IM conversation may have been too far apart to appear simultaneously within participants' field of vision, given the large 27 inch screens which were used at high resolution. Therefore, users would have to make the decision to look at the IM window at random or estimated intervals and not in response to a signal from the IM client. This would not rule out the first explanation entirely, since real-time text could still have grabbed participants' attention during those gazes and forced a decision from the user about whether to read an entire message or return to the Anticipation game. Further insight into this issue could be attained by replicating the experiment using eye-tracking techniques. An analysis of user eye-movement could be helpful in determining whether the motion involved in real-time text draws users' eyes away from a present task.

It is also possible that even if both windows were available within users' field of vision simultaneously, the text in the IM window may not have been salient enough to draw participants' attention away from the engaging game. The default settings of the AIM client present the text at the bottom of the window and in a relatively small font. It may simply have been hard to notice that text was incoming without explicitly looking for it. Replicating this experiment using significantly more salient text could lead to a better explanation of these results.

Another experimental factor which makes these data less than conclusive is that the Anticipation game may have been a more engaging task than other activities frequently performed on computers. The results may have been different had participants been doing something very boring while being interrupted by incoming messages. Since boredom at waiting for a message has been reported as a usability challenge to IM and chat (Voida et al., 2002;

Vronay et al., 1999), it could be useful to investigate whether users would find real-time text less boring than some sort of uninteresting task.

Despite the influences of the experimental design, this experiment has at the very least shown that real-time text can be implemented in a way which is not significantly distracting from an engaging task. This suggests that games which implement IM or chat features can strongly consider adopting real-time text without concern that it will draw attention away from other aspects of the game.

Conclusion

Research goals revisited

This research had three primary goals.

1. Investigate the nature of communication through real-time text and compare it to spoken dialogue and to message-by-message IM. Determine whether or not it truly achieves a more accurate simulation of spoken dialogue than message-by-message IM.
2. Determine the effect of this simulation on the overall usability of IM.
3. Propose design considerations for collaborative tools from the results of these analyses.

Written or spoken communication?

The data from these studies provides some support that real-time text does achieve a closer approximation to spoken dialogue than message-by-message IM in at least a few specific ways. Users do not disrupt adjacency pairs as frequently when using real-time text. They spend less time “talking” simultaneously. There was evidence that real-time text can support collaborative completion just as spoken dialogue does. There was also evidence that once an utterance had been made, users were less likely to edit it than they were when using message-by-message IM. Since editing of an utterance is not possible in spoken dialogue, this finding suggests that users were more closely obeying the conventions of spoken dialogue than written dialogue.

However, it is also clear that the ambiguity between written and spoken communication described by Voida et al. (2000) is not completely resolved by real-time text. Users still disrupted adjacency pairs at a fairly high rate (29%), and the 10.8% overlap in the turn space (when users would be typing simultaneously) is twice what has been documented in face-to-face conversation (Levinson, 1983), and still higher than the 8% overlap which has been observed in phone conversations (Shriberg et al., 2001). And although it was clear that collaborative completion was *supported*, it was also clear that users avoided explicitly engaging in this behavior, despite their recognition of the possibility and their implicit collaborative completion (when they would mentally predict how incoming messages would end). There were experimental factors (i.e. pairing strangers rather than acquaintances) which may have contributed to this, but is also likely that an interface could do more to support collaborative completion if it was decided that supporting the closest possible approximation of spoken dialogue was an important goal of the interface. Likewise, although users self-edited their messages much less frequently than in message-by-message IM, they still exhibited this behavior. Because this is not possible in spoken dialogue, it can be said that a perfect simulation is not achieved. Overall, real-time text has added some clarity to the somewhat ambiguous question (Voida et al., 2002) of which types of conventions are obeyed in IM. Real-time text brings IM closer to aligning with the conventions of verbal communication, but it still maintains several of the characteristics of written communication which lead to that ambiguity.

Usability

This research has uncovered a complex relationship between real-time text and the usability of an IM tool. While it is apparent that real-time text, when examined in isolation as in each of these studies, does not result in universally *better* usability for an IM client, it does change the user experience in many ways. It is also apparent that a number of factors influence the usability of real-time text, many of which are context specific. There was evidence that the nature of the discourse influences the usability of real-time text. For instance, the usability of the IM client for conversations in which one user is doing all the talking and the other just listening is impacted by real-time text. And within that context, the usability of real-time text is different for the person reading messages than for the person typing them. The experience level with IM is also a factor which contributes to usability.

Some factors which do not seem to influence the usability of real-time text are multitasking and the typing ability of users. This is valuable because it frees system designers to ignore these factors when making design choices and to focus on more relevant factors.

Design implications

- **Reading versus writing.** Users have different needs when they are reading a message than when they are writing a message. Real-time text supports the message recipient at the expense of the message producer. Future designs could look into ways to support the reader's need to see new information quickly and consistently without infringing on the writer's freedom to edit a message as they produce it. Precisely measured delays of a few seconds could be implemented which give the writers time to make quick edits or fix typos before the recipient sees the message. Real-time text could also be an option which itself can be rapidly controlled in real-time at the front of the interface as opposed to a setting which must be accessed at another level of the interface. For example, by default the interface could mimic message-by-message IM, but at any point users could make some extra effort, such as holding down the mouse button, to allow them to "peek" at what the other person is typing. Another potential strategy could be to make the interface unbalanced within a conversation, letting some users see in real-time and others in traditional form.
- **Turn-taking versus editing.** There appears to be a tradeoff between supporting turn-taking versus supporting users' freedom to edit their own messages. Real-time text supports better turn-taking at the expense of editing support. System designers need to

take this tradeoff into consideration. Systems which want to support better turn-taking, such as systems designed for dialogue and a simulation of spoken dialogue (i.e. text telephones or virtual meeting tools) would seem to want to support turn-taking, and in these contexts real-time text would be valuable. However, some systems may not want to suppress users from feeling that they can openly edit their contributions. For instance, the productivity of collaborative document editing tools could be compromised if users feel less inclined to change something because their collaborators have seen them typing it. Also to be taken into consideration should be Herring's (1999) observation that the apparent interactional incoherence of CMC, and the turn-taking problem in particular, can lead to a more playful and at times intense form of communication. Herring argues that "the availability of a persistent textual record of the conversation renders the interaction cognitively manageable, hence offsetting the major 'negative' effect of incoherence in spoken interaction." She also demonstrates that users inherently recognize that the rules of turn-taking cannot be perfectly applied in CMC, but that this can actually be liberating and open up new channels for wordplay and humor, as well more opportunities for hyperpersonal communication (holding multiple conversations with multiple people simultaneously). The inclusion of real-time text may "re-impose" the restrictions of coordinated turn-taking to some extent, and designers should consider whether the liberation of these rules may be beneficial to their system.

- **Problem solving.** The finding that problem-solving strategy may have been influenced by real-time text is noteworthy and should be considered when designing collaborative tools. The context in which a system will be used and the goals for interaction it wants to support are relevant to a decision regarding the implementation of real-time text.

This research offers support for a context aware design which may alter elements of the interface, including the style of text display, depending on the specific goals of the interaction. Significant qualitative analysis is required on these data and on future data sets in order to more specifically characterize the different strategies, if they indeed exist, which are supported by real-time text or by message-by-message IM.

- **IM expertise.** Since the usability of real-time text was influenced by users' typical IM habits, there may be an expertise effect which should be considered. Because heavy IM users liked message-by-message IM, it is possible that these users have developed strategies to compensate for the things which real-time text seemed to support better, such as turn-taking. Herring (1999) proposes that there are several potential benefits to a lack of interactional coherence in CMC, many of which stem from users developing new strategies to work around these problems. Therefore, the expertise of the user is an important design consideration. Designers should consider whether they are developing for users who are regular IM users or for novices, and should consider how to support the needs of both. Real-time text certainly seems to offer benefits to novice users, although it is possible that an introduction to IM with real-time text may prevent users from learning the strategies for benefiting from the playful interactional incoherence Herring describes, as well as any other strategies or beneficial behaviors which derive from the style of interaction facilitated by message-by-message IM clients.
- **Multitasking.** Because real-time text did not influence users' ability to multitask in this experiment, designers should be aware that real-time text alone may not be a salient distraction. Systems such as QnA (Avrahami & Hudson, 2004) which want to afford users freedom to control the saliency of their messages on the recipients' screens, will not achieve this by simply implementing real-time text.

- **Collaborative completion.** Real-time text supports the possibility of collaborative completion, but this study found that there were other barriers, such as etiquette, which prevented users from explicitly finishing each others' utterances. As collaborative completion is an important characteristic of spoken conversation (Lerner, 2004) and has been shown to benefit collaborative learning and problem solving (Roschelle & Teasley, 1992), systems should consider ways to facilitate this interaction. Real-time text alone seems to be insufficient, and designers could look at ways of supplementing it through a shared space for message composition.

Future work

This study has presented a number of important directions for future research. A qualitative analysis of the differences in the style of dialogue and the problem solving strategies resulting from real-time text could lead to a better contextualization of its implementation. This research has also identified several factors which were not explicitly tested but which may influence the usability of real-time text. The familiarity of participants with each other could impact how they are influenced by real-time text, and an important future variation of these experiments would use partners who are closely familiar with each other, such as spouses, roommates, close friends, or siblings. Previous experience communicating via IM with each other would also be an important factor of that research. Other factors which future research on this topic should consider as potential independent variables are the baseline typing ability of participants, the goal of the interaction, surrounding interface elements, the environment in which IM is being used, and collaborative or communicative systems other than IM.

Why does this matter?

The usability of real-time text has important implications both to the design of IM systems and to any other system of synchronous, text-based communication and collaboration. Important workplace tools such as virtual meeting systems or collaborative document editing tools can be influenced by real-time text, and this research has presented some important considerations to their design. Real-time text is being implemented into text telephones for the hearing impaired, and the usability of these devices is crucial to the accessibility of telecommunication. It is particularly important to point out that the hearing impaired need to be able to communicate with everyone, and so the implementation of synchronous, text-based telecommunication will ultimately need to extend beyond specialized phones and into mainstream devices. Similarly, IM is becoming a standard feature in smartphones, and therefore this research has implications for the design of mobile interfaces. Companies are frequently adopting online chat as an option for customer or technical support. Efficient support systems can save these companies money, and this research has presented several variables which can be considered to maximize the efficiency of these types of support systems. The design of games which depend on collaboration can be influenced by an intelligent implementation of a synchronous text-based tool. And there are undoubtedly numerous other contexts in which the efficiency and usability of a communication tool can be improved by the data presented here. Real-time text has already been deployed in some communications systems, and this research has provided empirical evidence which can be used to optimize any future implementations.

Bibliography

About Google Wave. (n.d.). . Retrieved March 29, 2010, from http://wave.google.com/about.html

Avrahami, D., & Hudson, S. E. (2004). QnA: augmenting an instant messaging client to balance user responsiveness and performance. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (pp. 515-518). Chicago, Illinois, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=1031607.1031692&coll=GUIDE&dl=GUIDE&CFID=83884973&CFTOKEN=19111502

Baron, N. S. (2004). See you Online: Gender Issues in College Student Use of Instant Messaging. *Journal of Language and Social Psychology*, *23*(4), 397-423. doi:10.1177/0261927X04269585

Baron, N. S. (2005). Instant messaging and the future of language. *Commun. ACM*, *48*(7), 29-31.

Berglund, T. Ö. (2009). Disrupted Turn Adjacency and Coherence Maintenance in Instant Messaging Conversations — Language@Internet. Retrieved January 4, 2010, from http://www.languageatinternet.de/articles/2009/2106

Campbell, J. D. (2004). Instant messages: a framework for reading between the lines. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (pp. 519-522). Chicago, Illinois, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=1031607.1031693&coll=GUIDE&dl=GUIDE&CFID=75544600&CFTOKEN=55411312

Campbell, J. D. (2005). Does spelling matter in instant messaging?: answers from measuring error correction frequency. In *CHI '05 extended abstracts on Human factors in* *computing systems* (pp. 1244-1247). Portland, OR, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=1056808.1056887&coll=GUIDE&dl=GUIDE&CFID=75544600&CFTOKEN=55411312

Clark, H., & Brennan, S. (1991). Grounding in communication. In *Perspectives on socially shared cognition* (pp. 127-149). American Psychological Association.

Clark, H. H., & Schaefer, E. F. (1989). Contributing to Discourse. *Cognitive Science*, *13*(2), 259-294. doi:10.1207/s15516709cog1302_7

Dabbish, L., & Kraut, R. E. (2004). Controlling interruptions: awareness displays and social motivation for coordination. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work* (pp. 182-191). Chicago, Illinois, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=1031607.1031638

Dillon, A. (2001). Beyond usability: process, outcome and affect in human-computer interactions. *Canadian Journal of Library and Information Science*, *26*(4), 57-69.

Garcia, A. C., & Jacobs, J. B. (1999). The Eyes of the Beholder: Understanding the Turn-Taking System in Quasi-Synchronous Computer-Mediated Communication. *Research on Language & Social Interaction*, *32*(4), 337. doi:10.1207/S15327973rls3204_2

Grinter, R. E., & Palen, L. (2002). Instant messaging in teen life. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work* (pp. 21-30). New Orleans, Louisiana, USA: ACM. Retrieved from http://scholar.google.com.proxy.lib.umich.edu/citation.cfm?id=587082

Hancock, J. T., & Dunham, P. J. (2001). Language Use in Computer-Mediated Communication: The Role of Coordination Devices. *Discourse Processes*, *31*(1), 91. doi:10.1207/S15326950dp3101_4

Herbsleb, J. D., Atkins, D. L., Boyer, D. G., Handel, M., & Finholt, T. A. (2002). Introducing instant messaging and chat in the workplace. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* (pp. 171-178). Minneapolis, Minnesota, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=503408&dl=GUIDE&coll=GUIDE&CFID=80866997&CFTOKEN=89541859

Herring, S. (1999). Interactional Coherence in CMC. *Journal of Computer-Mediated Communication*. Retrieved from http://jcmc.indiana.edu/vol4/issue4/herring.html

Hillstrom, A. P. (1994). Visual motion and attentional capture. *Perception and Psychophysics*, *55*(4), 399.

Isaacs, E., Walendowski, A., Whittaker, S., Schiano, D. J., & Kamm, C. (2002). The character, functions, and styles of instant messaging in the workplace. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work* (pp. 11-20). New Orleans, Louisiana, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=587078.587081

Lenhart, A., Hitlin, P., & Madden, M. (2005). Teens and Technology | Pew Research Center's Internet & American Life Project. Retrieved March 8, 2010, from http://www.pewinternet.org/Reports/2005/Teens-and-Technology.aspx

Lenhart, A., & Shiu, E. (2004). 53 million adults trade instant messages and 24% of them swap IMs more frequently than email. IM also gains a following in U.S. workplaces | Pew Internet & American Life Project. Retrieved September 17, 2009, from http://www.pewinternet.org/Reports/2004/How-Americans-Use-Instant-Messaging.aspx?r=1

Lerner, G. H. (2004). *Conversation analysis*. John Benjamins Publishing Company.

Levinson, S. C. (1983). *Pragmatics*. Cambridge textbooks in linguistics. Cambridge

[Cambridgeshire]: Cambridge University Press.

Lewis, J. R. (1995). IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *Int. J. Hum.-Comput. Interact.*, *7*(1), 57-78.

McCarthy, J. C., & Monk, A. F. (1994). Measuring the quality of computer-mediated communication. *Behaviour & Information Technology*, *13*(5), 311. doi:10.1080/01449299408914611

Nardi, B. A., Whittaker, S., & Bradner, E. (2000). Interaction and outeraction: instant messaging in action. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work* (pp. 79-88). Philadelphia, Pennsylvania, United States: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=358975&dl=GUIDE&coll=GUIDE&CFID=80866997&CFTOKEN=89541859

Nasah, A. M. L. (2008). *Examining instant messaging impact on learning using an integrated worked-example format*. University of Central Florida. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=1467366&coll=GUIDE&dl=GUIDE&CFID=80866997&CFTOKEN=89541859

Neill, D. B. (2003). Cooperation and coordination in the turn-taking dilemma. In *Proceedings of the 9th conference on Theoretical aspects of rationality and knowledge* (pp. 231-244). Univerity of Indiana, Indiana: ACM. doi:10.1145/846241.846270

O'Neill, J., & Martin, D. (2003). Text chat in action. In *Proceedings of the 2003 international ACM SIGGROUP conference on Supporting group work* (pp. 40-49). Sanibel Island, Florida, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=958160.958167&coll=GUIDE&dl=GUIDE&CFID=80866997&CFTOKEN=89541859

Phillips, B. (2000). Should we take turns?: a test of CMC turn-taking formats. In *CHI '00 extended abstracts on Human factors in computing systems* (pp. 341-342). The Hague, The Netherlands: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=633292.633497&coll=ACM&dl=ACM&CFID=51877796&CFTOKEN=57948654

Real-Time Text Taskforce. (n.d.). . Retrieved January 4, 2010, from http://www.realtimetext.org/index.php

Real-time text versus other text based communications. (n.d.). . Retrieved September 14, 2009, from http://www.ictrnid.org.uk/whyrtt.html

Roschelle, J., & Teasley, S. (1992). *Journal of the Learning Sciences*, *2*, 235-276.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, *50*(4), 696-735.

Scupelli, P., Fussell, S. R., Kiesler, S., Quinones, P., & Kusbit, G. (2007). Juggling Work Among Multiple Projects and Partner. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences* (p. 77). IEEE Computer Society. Retrieved from http://scholar.google.com.proxy.lib.umich.edu/citation.cfm?id=1255480.1255634&coll=GUIDE&dl=GUIDE&CFID=72504546&CFTOKEN=83578991

Shriberg, E., Stolcke, A., & Baron, D. (2001). Observations on Overlap: Findings and Implications for Automatic Processing of Multi-Party Conversation. *EUROSPEECH-2001*, 1359-1362.

Soukoreff, R. W., & MacKenzie, I. S. (2003). Metrics for text entry research: an evaluation of MSD and KSPC, and a new unified error metric. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 113-120). Ft. Lauderdale, Florida, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=642611.642632&type=series

Voida, A., Newstetter, W. C., & Mynatt, E. D. (2002). When conventions collide: the tensions of instant messaging attributed. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* (pp. 187-194). Minneapolis, Minnesota, USA: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=503410&dl=ACM&coll=ACM&CFID=45023728&CFTOKEN=99511593

Vronay, D., Smith, M., & Drucker, S. (1999). Alternative interfaces for chat. In *Proceedings of the 12th annual ACM symposium on User interface software and technology* (pp. 19-26). Asheville, North Carolina, United States: ACM. Retrieved from http://portal.acm.org.proxy.lib.umich.edu/citation.cfm?id=320719.322579&coll=ACM&dl=ACM&CFID=45023728&CFTOKEN=99511593

Wittenbaum, G. M., Hollingshead, A. B., & Botero, I. C. (2004). From cooperative to motivated information sharing in groups: moving beyond the hidden profile paradigm. *Communication Monographs*, *71*(3), 286. doi:10.1080/0363452042000299894

Appendices

Appendix A: Task sheet for movie task

You and your partner are organizing a 3 day weekend retreat for 10 people to a cabin.

First task: Choose 5 movies that you want to rent to bring to the cabin

When you have completed this task answer the following question: Do not answer if you did not finish the task.

On a scale of 1 to 10, how satisfied are you with the movies you agreed to?
1 - Very Unsatisfied
10 - Very Satisfied

Your satisfaction? ______

Second task: You need to make a shopping list for the trip. You are trying to spend less than $200 on supplies for the trip. Decide on the items you will need to buy, using your best estimations of what each item will cost to keep it within the budget.

When you have completed this task answer the following question: Do not answer if you did not finish the task.

On a scale of 1 to 10, how satisfied are you with the movies you agreed to?
1 - Very Unsatisfied
10 - Very Satisfied

Your satisfaction? ______

Appendix B: Task sheet for transportation task

You and your partner are being asked to speak to University administrators on behalf of the students at U-M about the school's transportation system, including busses, walkways, and parking. You will have only 5 minutes to present to the administrators.

First task: With your partner, determine the three most important issues you want to bring to the Administration's attention.

When you have completed this task answer the following question: Do not answer if you did not finish the task.

On a scale of 1 to 10, how satisfied are you with your decision
1 - Very Unsatisfied
10 - Very Satisfied

Your satisfaction? ______

Second task: With your partner, develop a plan for gathering the required information for this study, such that you feel you can adequately speak for the student body. Determine what information will be required and how you will obtain it. Then, discuss how you would combine all this information to be able to make a strong argument in the limited amount of time.

When you have completed this task answer the following question: Do not answer if you did not finish the task.

On a scale of 1 to 10, how satisfied are you with your decision
1 - Very Unsatisfied
10 - Very Satisfied

Your satisfaction? ______

Appendix C: Computer System Usability Questionnaire

(Questions 1-19 answered on a 7 point Likert scale ranging from "Strongly disagree" to "Strongly agree.")

1. Overall, I am satisfied with how easy it is to use this system
2. It was simple to use this system
3. I can effectively complete my work using this system
4. I am able to complete my work quickly using this system
5. I am able to efficiently complete my work using this system
6. I feel comfortable using this system
7. It was easy to learn to use this system
8. I believe I became productive quickly using this system
9. The system gives error messages that clearly tell me how to fix problems
10. Whenever I make a mistake using the system, I recover easily and quickly
11. The information (such as online help, on-screen messages, and other documentation) provided with this system is clear
12. It is easy to find the information I needed
13. The information provided for the system is easy to understand
14. The information is effective in helping me complete the tasks and scenarios
15. The organization of information on the system screens is clear
16. The interface of this system is pleasant
17. I like using the interface of this system
18. This system has all the functions and capabilities I expect it to have
19. Overall, I am satisfied with this system

List the most **negative** aspect(s):

List the most **positive** aspect(s):

Appendix D: Murder Mystery Scenario for FBI Agent

You are an FBI agent and have been working to solve the murder of Mr. Body. You have spent all day collecting evidence, and you are now going to meet with a detective from the Chesterfield Police Department who also spent the day collecting evidence. You are to work together and come to a conclusion as to who should be arrested for the murder. The detective may have evidence which you have not collected and which could be useful to you, and you may have some evidence which the detective does not have.

There are 4 suspects in the case, Mrs. Peacock, Colonel Mustard, and Professor Plum and Miss Scarlet. You are to make an arrest according to 3 equally important criteria:

1. Access to the murder weapon
2. Lack of a credible alibi. An alibi is proof that the suspect was not at the scene of the crime when the crime occurred.
3. A motive for committing the crime. You do not need undisputable proof of a motive, but you should have some sort of reasonable suspicion based on the evidence.

Your task is to make the best decision as quickly as possible. Submit using the web form on your screen.

Below are the details of the murder and the evidence you have gathered:

The murder took place on Thursday March 12 in the Driscoll Building in downtown Chesterfield, where Mr. Body worked. Surveillance cameras show a masked assailant walk into Mr. Body's office and leave a few minutes later. A janitor on the floor below heard shots and later found Mr. Body dead in his office. You were just informed that the bullets which killed Mr. Body were .22 caliber.

Here is the information you have about each suspect:

<u>Mrs. Peacock</u>
A friend of Mrs. Peacock's told you that at a party recently, Mrs. Peacock received a text message and suddenly became outraged, claiming that an enormous sum of money had been stolen from her .
On the night of the murder, her housekeeper, who works from 1 pm to 9 pm, spoke with Mrs. Peacock in the house right before she finished working.

<u>Colonel Mustard</u>
His son claims that Colonel Mustard owns several pistols and two hunting rifles.
His alibi is that he was eating dinner at a restaurant down the street from the Driscoll Building when the murder occurred.
The restaurant's credit card records confirmed that he ate dinner there that night, and that he paid at 8:45 pm.

<u>Professor Plum</u>
Gun records show that Professor Plum owns a .22 caliber pistol and a shotgun.

His alibi is that he was teaching when the murder occurred.
Professor Plum teaches an evening class on Tuesdays and Thursdays across the street from the Driscoll Building.
His students confirmed that he was teaching, but also that he let class out 30 minutes early each day during the week of the murder.
Professor Plum normally holds office hours just prior to class, but this week students reported he cancelled all his office hours.

Miss Scarlet
Miss Scarlet's boyfriend was arrested on March 13 for illegal possession of firearms. Witnesses saw him put a pistol in his gym bag at school and called the police.
Miss Scarlet claims she was in Professor Plum's class when the murder took place on the evening of March 12th.
An email between Miss Scarlet and Mr. Body on March 9th suggests that Mr. Body had uncovered a hidden bank account which was collecting money from an overseas investment and to which Miss Scarlet had access.
Miss Scarlet rents a room from Mrs. Peacock. The security guard at the gate to the community recorded that Miss Scarlet returned on March 12th at 10:15 pm.

Appendix E: Murder mystery scenario for police detective

You are a detective for the Chesterfield Police and have been working to solve the murder of Mr. Body. You have spent all day collecting evidence, and you are now going to meet with an FBI agent who has also spent the day collecting evidence. You are to work together and come to a conclusion as to who should be arrested for the murder. The FBI agent may have some evidence which you have not collected and which could be useful, and you may have some evidence which the FBI agent does not have.

There are 4 suspects in the case, Mrs. Peacock, Colonel Mustard, and Professor Plum, and Miss Scarlet. You are to make an arrest according to 3 equally important criteria:

1. Access to the murder weapon
2. Lack of a credible alibi. An alibi is the suspect's statement and supporting evidence about where they were when the crime took place.
3. A motive for committing the crime. You do not need undisputable proof of a motive, but you should have some sort of reasonable suspicion based on the evidence.

Your task is to make the best decision as quickly as possible. Submit using the web form on your screen.

Below are the details of the murder and the evidence you have gathered:

The murder took place on March 12 in the Driscoll Building in downtown Chesterfield, where Mr. Body worked. Surveillance cameras show a masked assailant walk into Mr. Body's office at precisely 9:30 pm carrying a pistol, and leave a few minutes later. A janitor on the floor below heard shots at that time and 20 minutes later found Mr. Body dead in his office. Mr. Body was an accountant who worked for each of the 4 suspects and had access to each of their detailed financial information.

Here is the information you have about each suspect:

<u>Mrs. Peacock</u>
You checked records and determined that Mrs. Peacock's husband owned a .32 caliber pistol. Mrs. Peacock's alibi is that she was home when the murder occurred. Mrs. Peacock lives an hour outside downtown Chesterfield.

<u>Colonel Mustard</u>
Colonel Mustard is also a suspect in a money laundering scheme.
A clerk at a firearms store claims that Colonel Mustard purchased .22 and .32 caliber bullets about a week before the murder.
His alibi is that he was eating dinner at a restaurant down the street when the murder occurred. Witnesses confirmed they saw him there that evening.

<u>Professor Plum</u>
Gun records show that Professor Plum owns a .22 caliber pistol and a shotgun.
His alibi is that he was teaching when the murder occurred.

Professor Plum teaches an evening class from 8 to 9:30 pm across the street from the Driscoll Building.
Professor Plum normally holds office hours just prior to class, but this week students reported he cancelled all his office hours.

Miss Scarlet

Miss Scarlet is a student in Professor Plum's evening class. She has a perfect attendance record, including the week of the murder.
Miss Scarlet has been audited 3 times by the IRS and currently owes $90,000 in back taxes. She has been warned that she will be indicted on tax evasion should she have any more discrepancies. She claims she cannot repay any taxes until she regains a steady source of income.
Her boyfriend was arrested on March 13th for illegal possession of firearms. A policeman who searched his vehicle claims that boxes of .22 caliber bullets were found in the car.
Miss Scarlet has been renting a room from Mrs. Peacock for 3 months.

Appendix F: Post-test questionnaire from turn-taking study

1. Which type of text display did you prefer? The real-time display of the other person's typing, or the message-by-message display? Please explain.

2. Were you comfortable with the other person seeing you type your messages? Did that affect the way you participated in the conversation in any way? Please explain.

3. When using the real-time display, did you find yourself trying to predict what the entire message would say based on the first few words or characters?

4. If so, did you ever begin typing a response based on an incorrect assumption of the incoming message? Please describe any instances of this.