Running Head: Bridging Video and Instant Messaging

Bridging the Social Nature of Video and Instant Messaging?

Daniel Zinzow

Carnegie Mellon University

May, 2005

Introduction

We are all immersed in a social and cultural context. We do not respond to environment mechanically, but rather we are complex human beings that perceive, think about, and sometimes distort information from our environment. People form construals of their social environment, which is one's perception, comprehension, and interpretation one has of the social world. (Lewin, 1943)

This construal influences one's self, but the construal itself is influenced by three major ideas. The first is the social influence of others, whether it be direct influence (such as advertisements), mere presence of others, or imagined presence of others (Aronson, Wilson, & Akert, 2004). The second is one's need to maintain high self-esteem. That is, to see themselves as good, competent, and decent (Aronson, 1992a, 1998; Baumesiter, 1993; Harter, 1993; Kunda, 1990; Pyszczynski, Solomon, Greenberg, & Stewart-Fouts, 1995; Stone, 1998; Thibodeau & Aronson, 1992; Tice, 1993). Lastly, one's construal is influenced by one's need to form an accurate picture of the world (Aronson, Wilson, & Akert, 2004). Even though people try to bend the facts to make them look better, they do not completely distort reality for it would not be adaptive to live in a fantasy world. One can still not form a truly accurate perception of the world, though, since one does not have all the information needed for 100% accuracy and one's expectations about the world gets in the way of one's perceptions of the world.

As a result of us living in a social world, it does not matter what the purpose that is behind the development of Computer-Mediated Communication (CMC), it will still have an effect upon peoples' social behaviors. And actually seen with some CMC devices, the actual use becomes a social one, such as instant messaging, or only supports the transmission of affective information, such as video conferencing.

If one takes a look at the basics of grounding and conversational structure, it can clearly be seen that instant messaging has many incoherent qualities to it, yet it is widely used for social purposes. Due to the incoherent quality and the lack of ability to transmit nonverbal cues, though, affective information is very difficult to send across the medium. On the other hand, video is a very rich media that can transmit nonverbal cues, and therefore has shown to support the transmission of affective cues. At first glance, it

seems as though the incoherence issues found with instant messaging could be resolved with the addition of video.

To test this though, it is necessary to take a deeper look into conversation structure and compare how instant messaging and video support or disrupt this structure. Also taking a look into what makes instant messaging popular and what makes video good for transmitting social cues will give a better idea of how these two mediums could be combined.

Conversational Structure

People need to have common ground before they can even begin to coordinate on content. Common ground being mutual knowledge, mutual beliefs, and mutual assumptions (Clark & Carlson, 1982; Clark & Marshall, 1981; Lewis, 1969; Schelling, 1960). Common ground is updated with a process called "grounding", in which participants try to establish that what was said has been understood (Clark & Shaefer, 1987, 1989; Clark & Wilkes-Gibbs, 1986; Isaacs & Clark, 1987). Once it is understood, it becomes part of their common ground.

Clark and Brennan (1991) listed eight constraints that a medium may impose on the communication between two people (A and B).

I . **Copresence***:* A and B share the same physical environment.

2. **Visibility***:* A and B are visible to each other.

3. **Audibility***:* A and B communicate by speaking.

4. **Cotemporality***:* B receives at roughly the same time as A produces.

5. **Simultaneity***:* A and B can send and receive at once and simultaneously.

6. **Sequentiality***:* A's and B's turns cannot get out of sequence.

7. **Reviewability***:* B can review A's messages.

8. **Revisability***:* A can revise messages for B.

The table below shows what constraints several types of media have.

Table 1. Seven Media and Their Associated Constraints (Clark & Brennan, 1991)

Table 1

SEVEN MEDIA AND THEIR ASSOCIATED CONSTRAINTS

Medium	Constraints
Face-to-face	Copresence, visibility, audibility, cotemporality, simultaneity, sequentiality
Telephone	Audibility, cotemporality, simultaneity, sequentiality
Video teleconference	Visibility, audibility, cotemporality, simultaneity, sequentiality
Terminal teleconference	Cotemporality, sequentiality, reviewability
Answering machines	Audibility, reviewability
Electronic mail	Reviewability, revisability
Letters	Reviewability, revisability

People are forced to used alternative grounding methods when the medium they are using lacks one of these characteristics. This happens because the costs of grounding techniques change. Clark and Brennan (1991) list those costs as

1. **Formulation Costs:** The cost of formulating an utterance
2. **Production Costs:** The cost of producing an utterance
3. **Reception Costs:** The cost of receiving the message
4. **Understanding Costs:** The cost of understanding the message
5. **Start-up costs:** The cost of starting up a new discourse
6. **Delay Costs:** The cost of delaying an utterance in order to plan, revise, and execute it more carefully
7. **Asynchrony Costs:** The cost of timing utterances
8. **Speaker Change Costs:** The cost of changing speakers, as in turn taking
9. **Display Costs:** The cost of displaying objects to the recipient
10. **Fault Costs:** The cost of producing an utterance fault
11. **Repair Costs:** The cost of repairing faults in utterances

Conversations are remarkably highly structured. People take turns speaking, generally stay on topic, quickly repair misunderstandings, and much more. Even though prior to its occurrence the course of a conversation is unpredictable, it still has apparent structure and order to it.

If the structure is not kept in a conversation, though, then there are breakdowns, confusions, and misunderstanding; therefore, no common ground is being established. So obviously the structure of conversation is important. There are three parts of conversation that make it structured: adjacency pairs, turn-taking, and coherence.

Adjacency Pairs

In conversations, people's utterances are constrained by the utterances of others. The smallest structural unit that exhibits this constraint is known as an adjacency pair. Adjacency pairs have the following general features, according to Schegloff and Sacks (1973).

" 1. They consist of two utterances, a first-pair part and a second-pair part

2. The two utterances are spoken by different speakers

3. The utterances are paired so that a first-pair part must precede a second-pair part

4. The first-pair part constrains what can occur as a second-pair part

5. Given the first-pair part, the second-pair part becomes conditionally relevant."
(Holtgraves, 2002)

It is not necessary for the second-pair part to follow immediately after the first-pair part in an adjacency pair because insertion sequences (Schegloff, 1972) or side sequences (Jefferson, 1972) can occur between these pairs. These side issues remain relevant, though, to the original issue so as to maintain the expectation that the second-pair will be coming.

Adjacency pairs represent a template for the production and interpretation of talk; given the first-pair part, the second-pair part becomes conditionally relevant (Schegloff, 1968). If the expected second pair part of an adjacency pair is absent this will lead to inferences about the person who failed to provide this second-pair (Holtgraves, 2002).

Turn-Taking

Conversations are amazingly ordered; one speaker speaks at a time and there is little or no gaps or overlaps during speaker transitions. A model proposed by Sacks, Schegloff, and Jefferson (1974) accounts for turn-taking regularity and has the following two components:

1. Turn-constructional component – Specifies what constitutes as a turn; once a turn starts it becomes known what type of turn it is and when it will end, which is known as projectability

2. Turn allocation component – Techniques used for allocation of turns, whether it is speaker appointed or self appointed

Projectability is determined by the intonational and syntactic properties of the turn (Holtgraves, 2002).

When the turn-taking system breaks down, conversants will repair the system, displaying peoples' orientation to this turn-taking system. Conversational repair is part of the conversational system, which is a system that is designed to facilitate communication.

Coherence of Topic

Conversations cohere on one topic or a set of topics. There is no formal manner, though, in which conversational coherence has been defined, yet people highly agree when segmenting conversations into topics (Planalp & Tracy, 1980). Since conversations are generally unplanned, conversational coherence is recognized as being an emergent property. People have an idea of what they want to talk about before starting a conversation, but the way in which these things get talked about emerges as the conversations unfolds. Even though there is no formal rule for coherence, conversationalists are oriented toward a topic continuation rule. This rule appears in conversation when people position their topic introductions to connect with prior turns made by others. This allows a prior topic to serve as a resource for starting the new topic and therefore the new topic will be seen as being relevant to this prior topic (Schegloff & Sacks, 1973).

Instant Messaging - Incoherence

Instant messaging is a text-based tool that allows users to conduct conversations in near synchronicity over the internet. Common features of instant messaging are contact lists, away messages, customizable backgrounds, and emoticons.

Instant messaging formed from computer networks that were originally intended for data transmission, not for social interaction (Reid, 1991; Rheingold, 1993; Walther,1996). There are various limitations on group interactions in instant messaging, including high production, reception, speaker changes costs, and tensions between the verbal communication and written communication that instant messaging encompasses (Clark & Brennan, 1991; Herring, 1999). This gives rise to four properties, out of many others, of the medium that are obstacles to interaction management.

1. Reception Cost – Lack of simultaneous feedback
2. Speaker Cost – Disrupted turn adjacency
3. Production Cost – Perceived impoliteness
4. Tension between verbal and written communication

Lack of Simultaneous Feedback

Instant messaging is a "lean" medium (Daft & Lengel, 1984, 1986) which has fewer channels than face-to-face (FtF) interaction for transmitting messages. With instant messaging, users do not see or hear their interlocutors and as a result do not have access to non-verbal information of the other.

In one-way systems, messages are sent in their entirety as soon as the writer presses 'enter', rather than having the message appear as it is being produced. As a result, it is impossible for the addressee to respond while the message is being written. Research on conversational interaction with speech shows "that simultaneous feedback plays an important role in signaling listenership, timing turn-taking effectively, and maintaining continuous interaction" (Herring, 1999; McLaughlin, 1984). The absence of this feedback may result in overlap in turn sequences, discontinuity in turn-sequences, and a difficulty in tailoring messages to others.

Disrupted Turn Adjacency

The other problem with one-way systems is that messages are transmitted in the order that they are received by the system. With disrupted adjacency, users may have difficulty tracing sequential exchanges resulting in fragmented interaction.

There may be no temporal overlap with instant messaging but there is a lot of "overlap exchanges" (Herring, 1999). Users may become impatient when unable to tell whether the interlocutor is in the process of responding, and send a second message. This then could result in an incomplete or interleaved exchanges sequence (Condon & Cech, 1996; Marvin, 1995). The situation becomes worse with multiparty communication. Unrelated messages from other participants intervene between initiating and response messages. (Cherney, 1995; Lunsford, 1996; Murray, 1989].

Even without intervening messages, instant messaging does not adhere to the ideal that the speaker turn alternates in an orderly manner. There isn't necessarily a one-to-one correspondence between the initiator and responder, for there can be multiple responses directed at a single initiating message. Single messages can also be directed towards multiple initiating messages. Or messages may even not receive any responses. When no response occurs, though, participants send out multiple messages to attract a response, which is a display of the difficulty users have in determining others' intentions due to the lack of feedback in instant messaging. In FtF conversations, if the second-pair in an adjacency pair is absent, then inferences are made (Holtgraves, 2002), but in instant messaging, it cannot be assumed that if one receives no response to a message that this is intentional, for there are many reasons the receiver did not respond. There could be technical problems, or they may not have noticed, or were absent momentarily.

With FtF conversations, there may be side sequences in adjacency pairs, but these side sequences stay relevant to the topic. With instant messaging, though, adjacency pairs are often disrupted by intervening, irrelevant messages. This results in quick decay in topics and off-topic digressions.

Perceived Impoliteness

It has been noticed by many that conversations conducted through instant messaging often seem less polite and less inhibited than FtF communication (Kiesler,

Zubrow, Moses, & Geller, 1985). This is mainly due to the affordances of the instant messaging medium, and may not be the result of people being impolite themselves.

There is also the cost of production, where it takes much more effort to type than speak. As a result, hedging, which appears to signal a speaker's level of commitment to an utterance, happens less often in typed than in spoken utterances. When hedges modify a speaker's utterance, it is a display that they grant the addressee the license to reject or modify the utterance. This option serves both task-related and face-management needs. Brown and Levinson (1978) present hedges in terms of people's needs to save face or avoid threatening the face of others.

In an experiment done by Brennan and Ohaeri (1999) there was a significant positive correlation between word counts and hedge rates. Taking word counts to be a measurement of a person's overall aversion to typing, it can be seen that the harder it is for one to type (i.e. the less the word count would be) the less there are of hedges. This correlation reveals that the reason less hedges are used is because of production cost, not because of people behaving less politely due to depersonalization caused by this medium.

Since hedging is a way for people to display their alignment toward their own and their partners' messages and therefore "achieve true politeness" (Brennan & Ohaeri, 1999), then the absence of hedging may result in the appearance of impoliteness even though such is not intended.

Tensions Between Verbal and Written Communication

Instant messaging is nearly synchronous, which is a characteristic shared with most verbal communication, but also allows users to attend to it when desired, which is a characteristic of most written communications. Voida, Newstetter, and Mynatt (2002) conducted an experiment in which participants implied that instant messaging is valued because of the "unique balance it holds in affordances between the conventions of verbal and written communication." The problem with this is that these two types of

communications are rather different, therefore causing a tension between the two within instant messaging. Voida, Newstetter, and Mynatt present five different types of tensions that occur in instant messaging.

1. **Persistence and Articulateness Tensions**

Users treat conversations casually and informally as with verbal communication, not worrying about editing their language. But with the persistence of the text on the screen, when errors occurred in spelling or grammar, the speaker made a fix quickly so not to be seen as illiterate as the typing may indicate.

2. **Synchronity Tensions**

This tension between the near-synchronicity of instant messaging and users' desire to make the interaction feel asynchronous. This causes users to get bored and start up multiple conversations, which results in difficulty in following all conversations going on. This can cause users to miss important information.

3. **Turn-Taking and Syntax Tension**

The exchange of the communicative artifact is not reliable enough to structure turn-taking, as in written communication, because both users are able to contribute at the same time. Also listeners cannot make gestures to indicate they wanted to speak, as in verbal communication.

4. **Attention and Context Tensions**

This tension is between the amount of attentiveness appropriate for instant messaging, with a lot of attention needed for verbal communication and little for written communication. Participants explained liking instant messaging because they felt they did not have to attend and respond right away; yet there was a need for users to justify their absence in responsiveness to the initiator. This justification provided context that would have been apparent in a verbal communication due to copresence.

5. **Availability and Context Tensions**

Tension between the nature of verbal communication, in which body language is an indicator of availability, and written communication, in which the initiator of the communication has no influence on when the communication will

be dealt with. There is conflict since the initiator doesn't know if the recipient is there, even if they actually are.

Instant Messaging – User Adaptations

As stated previously, "if the structure is not kept in a conversation, then there are breakdowns, confusions, and misunderstanding; therefore, no common ground is being established. So obviously the structure of conversation is important. There are three parts of conversation that make it structured: adjacency pairs, turn-taking, and coherence." As can be seen from the above explanations, instant messaging is hardly structured at all.

As Clark and Brennan (1991) state, when a medium is lacking one of the eight characteristics of mediums, people are forced to use alternative grounding methods to create common ground due to the costs of grounding imposed by a medium. Instant messaging has reception, production, and speaker costs, so there must me some technique that instant messaging has to respond to these costs. What these techniques are, are still a question. Some ideas though are that users of instant messaging have devised and use a set of programming "shortcuts" to type the most common of them (such as giggle, happy, sad). An example of this is what is commonly known as 'emoticons.' Other practices are to use symbols at the end of a message to indicate that one is not yet ready to give up the floor or 'raising one's hand' by using a command to represent (Herring, 1999).

A method some users use to deal with intervening messages and multiple participants is to preface a turn with the user name of the intended addressee (Werry, 1996).

Instant Messaging – Popularity and Usage

(Following statistical information is from Lenhart & Shiu, 2004)

73% of all adults go online, but this percentage is higher if those above 60 are not included. The breakdown by generation is:

87 % of teens go online

84% of Generation Y (18-28) go online

87% of Generation X (29-40) go online

79% of Trailing Boomers (41-50) go online

75% of Leading Boomers (51-60) go online

54% for Matures

and 21% for the 70 plus in age.

There are many things to do online, so the actual percentages for use of instant message is much lower, with 47% of adults using instant messaging and a range from 25 – 75% within each generation group.

Table 2. Generational Differences in Online Activites

Generational Differences in Online Activities

	Online Teens[a] (12-17)	Gen Y (18-28)	Gen X (29-40)	Trailing Boomers (41-50)	Leading Boomers (51-59)	Matures (60-69)	After Work (70+)	All Online Adults[b]
Go online	87%	84%	87%	79%	75%	54%	21%	72%
Teens and Gen Y are more likely to engage in the following activities compared with older users:								
Online games	81	54	37	29	25	25	32	36
School research	*	73	60	61	48	33	14	57
Instant message	75	66	52	38	42	33	25	47
Text message	*	60	44	29	15	11	8	35
Get info about a school	57	59	42	50	40	30	14	45
Download music	51	45	28	16	14	8	5	25
Read blogs	38	41	30	20	21	19	16	27
Download video	31	27	22	14	8	8	1	18
Create a blog	19	20	9	3	9	3	4	9
Activities where Gen X users or older generations dominate:								
Get health info on at least one topic	*	73	84	80	84	68	72	79
Travel reservations	*	50	72	64	64	59	60	63
Job research	*	44	59	59	54	31	13	51
Use gov't sites	*	41	56	64	60	55	45	54
Bank online	*	38	50	44	37	35	22	41
Religious info	26	30	38	24	28	28	28	30
And for some activities, the youngest and oldest cohorts may differ, but there is less variation overall:								
Use email	89	88	92	90	94	90	89	91
Get news	76	72	76	75	70	74	68	73
Product research	*	79	80	83	79	74	60	78
Online purchase	43	68	69	68	67	65	41	67
Job hunting	30	62	51	40	36	17	2	44
Use a photo service	*	39	38	31	32	31	30	34
Rate a person or product	*	36	34	27	31	24	8	30
Search for a person	*	31	31	23	23	24	29	27
Participate in an online auction	*	26	29	25	20	18	6	24

[a] Source for Online Teens data: Pew Internet & American Life Project Teens and Parents Survey, Oct.-Nov. 2004. Margin of error is ±4% for online teens.

[b] Source for Online Adult data: Pew Internet & American Life Project Surveys conducted January 2005, May-June 2005, and September 2005. Margin of error for all online adults is ±3% for these surveys. Questions pertaining to health topics, government sites, and religious info were asked in November 2004. Margin of error for that survey is ±5% for all online adults. The average margin of error for each age group, which can be considerably higher than ±5%, is listed in the Methodology section of this report.

77% of Instant Messaging (IM) users reported using IM at home, 21% reported using IM at work, and 7% report using it at school.

47% of IM users say they actively IM for 15 minutes or less during their usual IM sessions with much smaller percentages using IM for a longer period of time than that.

34% percent of IM users have posted a profile for their IM screen name, which include information like contact information, important personal news, inspirational quotes, funny sayings, links to interesting web sites, and links to personal photos.

47% of IM users have blocked someone from sending them instant messages and 54% have ever removed a buddy from their buddy list.

21% of people surveyed used IM at work of these:

     40% indicated they generally IM coworkers

     33% reported instant messaging with friends and family

     21% responded that they IM with both groups equally.

     40% feel that IM has improved teamwork

     50% believe IM saves some to a lot of time

The table below shows that people do multi-task when using instant messaging.

Table 3. Young IM-ers Perform other Tasks while Instant Messaging

Younger IM-ers perform other tasks while instant messaging

Percentage of IM users	Total Users	Gen Y	Gen X	Trailing Boomers	Leading Boomers	Matures	After Work
Do other things on the computer or internet, such play games for browse the internet							
Every time or almost every time	32%	49%	32%	25%	20%	<1%	10%
Just some of the time	29	28	33	33	17	27	32
Never	38	23	33	41	62	67	58
Do other things NOT on the computer or internet, such as talk on the phone or watch TV							
Every time or almost every time	20%	31%	17%	15%	10%	10%	10%
Just some of the time	30	36	37	26	17	15	11
Never	49	34	46	56	71	69	75

Source: Pew Internet & American Life Project Tracking Survey, February 2004. Margin of error is ±5%.

All of these statistics show that instant messaging is rather prevalent in our society, being used at home, work, and even school. It seems that instant messaging is used more as a short communication method rather than having drawn out conversations for hours. And neither is it very intense since people are able to multi-task even with things that are not

on the computer. Instant messaging also allows users to display personal and interesting information for others to see with profiles of themselves, but they still have control over who sees what with the ability to block others and also the ability to delete 'buddies' from a list.

Video and Nonverbal Hypothesis

Communication requires the coordination of both content and process between speakers and listeners (Clark & Brennan, 1991; Whittaker, Brennan, & Clark, 1991). Content coordination is how participants build up common ground and process coordination is the mechanisms and management of conversations.

Whittaker and O'Conaill (1997) analyzed the types of visible information that are used to support content and process coordination. The following tables are the results of their research. Each table presents different mechanisms for process coordination and content coordination and how the visible behavior affects those areas.

Table 4. Communicative Functions of Gaze

	Gaze Behaviors		
	Conversational Mechanism	Speaker	Listener
Process Coordination	Turn-taking cues	Speaker predominantly looks away from listenr while talking – negotiated mutual gaze used as "turn-yielding" signal	Listener predominantly looks at speaker while speaker is talking: negotiation mutual gaze used as "turn-accepting" cue
	Availability Cues		
Content Coordination	Reference	Gaze at an object indicates person's interest and attention to that	Gaze at an object indicates person's interest and attention to that

		object: joint attention allows pointing	object: joint attention allows pointing
	Feedback Cues	Gaze at listener can be an attention-eliciting device	Gaze at the speaker indicates interest in what the speaker is saying
	Interpersonal Information Cues	Patters of gaze interpreted as indicating sincerity, trustworthiness, friendliness; indicate speakers' affective attitude to utterance	Patters of gaze interpreted as indicating sincerity, trustworthiness, friendliness; indicate listeners' affective attitude to utterance

Table 5. Communicative Functions of Gesture

		Gestural Behaviors	
	Conversational Mechanism	Speaker	Listener
Process Coordination	Turn-taking cues	Termination of speaker gesture interpreted as "turn-yielding" cue	Listener gestures signal desire to speak
	Availability Cues		
Content Coordination	Reference	Pointing facilitated by joint attention	Pointing facilitated by joint attention
	Feedback Cues		
	Interpersonal		

Information Cues

Table 6. Communicative Functions of Facial Expressions

		Facial Expression	
	Conversational Mechanism	Speaker	Listener
Process Coordination	Turn-taking cues		
	Availability Cues		
Content Coordination	Reference	Visual Information from reading the speaker's lips decreases the ambiguity of speech	
	Feedback Cues		Head nods indicate assent or dissent; expressions indicate interest, understanding, puzzlement, or disbelief
	Interpersonal Information Cues	Expressions indicate happiness, fear, interest, surprise, sadness	Expressions indicate happiness, fear, interest, surprise, sadness

Table 7. Communicative Functions of Posture

Gestural Behaviors			
	Conversational Mechanism	Speaker	Listener
Process Coordination	Turn-taking cues		Listener activity can signal a desire to interrupt
	Availability Cues		
Content Coordination	Reference		
	Feedback Cues		Attention, interest in what speaker is saying
	Interpersonal Information Cues	Reveals speaker's attitude to utterance	Reveals listener's affective reaction to an utterance

Whittaker and O'Conaill (1997) proposed a nonverbal communication hypothesis that states that visible behaviors such as the ones expressed above provide information that is absent from audio-only communication.

This hypothesis is broken down into three versions (Whittaker, 1995, 1996):

1. Video provides cognitive cues that facilitate shared understanding
2. Video offers process cues to support turn-taking
3. Video provides social cues and access to emotional information

Little impact of visual information on cognitive problem solving have been showed by several laboratory studies (Chapnis, 1975; Chapanis et al., 1972; Reid, 1977; Short et al., 1976; Williams, 1977). Even FtF interaction is no better than speech only communication for collaborative problem solving.

Sellen (1995) found little evidence to support the claim that high-quality video information improves turn-taking and conversation management. There was no process differences between the video/audio systems and speech-only communication. Neither

video or audio systems replicated FtF conversational processes. The video system did reduce listeners from spontaneously taking the floor and led speakers to use more formal techniques for handing over conversation. Video/audio is *perceived* to be better than speech. It is perceived to "support interruptions; lead to more natural conversations; increase the ability to listen selectively to particular speakers; allow one to determine whether one is being attended to; and to generally keep track of the conversations" (Whittaker & O'Conaill, 1997; Sellen, 1995).

Video did have an effect on transmitting social cues and access to emotional information. When video information is added to the speech channel, the outcome and character of communication tasks that require access to affect or emotional factors is changed. Example tasks include negotiation, bargaining, and conflict resolution. When participants have access to visual information, they focus more on the motives of others and their conversations are more personalized, less argumentative, more polite, broader in focus, and less likely to end in deadlock than in speech-only communications (Reid, 1977; Short et al., 1976; Williams, 1977). In addition, groups conversing using audio with video tend to like each other more (Reid, 1977; Short et al., 1976; Williams, 1977).

Video and Informal Communication

As explained above, many experiments have shown for video to not be beneficial for cognitive cuing or conversation management. So instead of focusing on that Fish, Kraut, Root, and Rice (1993) devised a video system to aid informal communication.

Both media richness theorists [6] and social presence theorists [24] place communication channels along a continuum with FtF interaction at the richer, social end and written documents at the other. Therefore, video teleconferencing should be well suited for informal communication, especially so for aiding the more social aspects of communication.

From Fish et al.'s experiments with the Cruiser system between mentors and students, many of the mentors reported that they used the system to inquire about project status but would meet FtF when commenting about the student's personal performance. One mentor mentioned that he scheduled a FtF evaluation session because he "anticipated a long meeting in which the parties would need a 'richer' communication environment,

including the ability to see and respond to subtle reactions" (Fish, Kraut, Root, & Rice, 1993). Respondents reported that the Cruiser conversations were less useful than FtF for learning about their conversational partner and they felt substantially more privacy violation. Another issue was that since cameras have only a fixed field of view, users were concerned that other people might be present at the other side.

So it seems that video does not support informal communication too well since it is not 'rich' enough and invades privacy. Even though video may support transmission of social cues and affect information, maybe it is not the best medium for doing such or at least not the ways it has been implemented so far.

What is Necessary of Video

The face is the primary carrier of emotion and therefore it is a major contributor to effective interpersonal communication and affective computing (Lisetti & Schiano, in press). As bandwidth increases, adding a video channel becomes possible, but still under constrained conditions. Typically, the tradeoff has been to display highly realistic facial images over motion, and therefore video has not been well accepted as an addition to audio channels.

The perception of emotional content relies heavily on motion information Bruce, 1996). Emotions can be identified in displays that consist only of fields of moving dots, which do not contain facial features (Bassili, 1979). The timing and trajectories of facial expressions are highly precise and people show great sensitivity to these temporal parameters in trying to determine felt emotion, its intensity, and sincerity (Colston & Schiano, 1995). In the study done by Ehrlich, Schiano, and Sheridan (2000), participants were presented with moving images which showed an actor transitioning from a "neutral" expression toward a very intense, specific emotion. Participants were also presented with static images which the endpoints from these transitions.

The moving images were rated higher for perceived intensity of the emotions than static images, despite the fact that on average lower-intensity expression were seen in the moving images condition. This average lower-intensity being due to the fact that the motion cues were derived from a transition between neutral and a specific emotion. Image degradation had no effect on moving images but did decrease perceived intensity

for static images. The motion cues effectively compensated for image degradation. Therefore, it seems that the right tradeoff to be made is motion over facial reality, at least for the sake of expressing emotions.

Video and IM: Together?

Video increases the transmission of social cues and affective information over audio and instant messaging has a very unstable structure to it, which limits its ability to support grounding and also the transmission of emotion. From the look of things, it seems that video would be perfect for instant messaging; it would provide the ability for users to send their emotions to other users without having to deal with the broken down structure of instant messaging. But one should not be so quick to make a decision that these two mediums could work together. First it is important to understand why instant messaging is so popular. As has been stated many times, its structure is very incoherent even with the user adaptations so there must be something about instant messaging to overcome this incoherency and make it popular.

Why is IM Popular?

There are several thoughts to why instant messaging is so popular. It is very interactive, it depersonalizes users, and it contains ambiguity.

Interactivity

Incoherence may be the norm in instant messaging. Relaxed norms can be liberating and allow the possibilities of language play. Danet, Ruedenberg-Wright & Rosenbaum-Tamari (1997) claim that CMC is "an inherently playful medium." The weakening of relevance norms invites humorous play and on-line surveys have even shown that humorous messages are the most highly appreciated types of messages in computer-mediated environments.

Another notable attraction is the ability to have multiple, simultaneous exchanges. As Herring (1999) puts it, CMC participation "maximizes stimulation and minimizes opportunities for boredom; that is, it reduces receiver cost." There is greater intensity of interaction with instant messaging than what is possible FtF. And with the persistence of

the text in instant messaging, users are able to participate in simultaneous multiple interactions without getting lost since as they can refer back to what was going on.

In an experiment done by Dubrovsky, Kiesler, & Siegel (1983), a simultaneous computer conversation program was devised that enforced procedural order on the conversations, so that it would increase its similarity to FtF communication. The program forced group members to take turns speaking and to indicate when they wished to interrupt. It only allowed one person to talk at a time. The most important outcome of this study was that the software that controlled the sequence of interaction was disliked, which supports the idea that participants in instant messaging enjoy the freedom from norms that it provides.

Depersonalization

Studies done by Siegel, Dubrovsky, Kiesler, and McGuire (1983) revealed that people in computer-mediated groups were more uninhibited than they were in FtF groups as measured by uninhibited verbal behavior, defined as frequency of remarks containing swearing, insults, name calling, and hostile comments. It was also found that group members using the computer participated more equally than they did when they talked fact to face. Numerous other studies show higher levels of uninhibited behavior in CMC, including increased intimacy and flaming (Kiesler, Siegel, & McGuire, 1984; Kiesler, & Sproull, 1992; Matheson & Zanna, 1990; Walther & Burgoon, 1992).

In a study done by Connell, Mendelsohn, Robins, and Canny (2001) participants interacted either by FtF, telephone, or computer chat. After having conversations with other participants, the participants were asked to rate how they perceived both their own and their partners' behavior intent. According to the self and partner reports, people acted more like themselves in both the telephone conversations and computer chats than in FtF conversations. This occurred because less rich media allows people to express themselves with less inhibition.

This study also focused on self-awareness. People can be both publicly and privately self-aware (Buss, 1980; Carver & Scheier, 1981). When publicly self-aware, people focus on how they appear to others, which often leads to a feeling of discomfort (Carver & Scheier, 1981). People become publicly self-aware when others are either

looking at them or ignoring them or if they are given feedback as to how they appear to others, such as in a mirror or video image of themselves (Davis & Franzoi, 1991). On the other hand, when people are self-aware, they experience an intensification of emotion and also a deeper understanding of their internal thoughts and feelings (Buss, 1980). People become privately self-aware when their attention is drawn inward to themselves, such as when writing in a journal or diary (Davis & Franzoi, 1991).

Ambiguity

Face-work is "the measures people take to preserve face for themselves and for others when problematic events occur during interactions" (Aoki & Woodruff, 2005). Social interaction often requires face-work since it is used to accomplish goals such as avoiding embarrassment and maintaining harmony in relationships. It also involves managing the impressions that other people have of your behavior. This can be done by making sure people have a correct understanding of your actions or giving a misleading impression by telling a 'white lie.'

An example given by Aoki and Woodruff (2005) is that when you try to contact a friend but get no response and later get an excuse, you may or may not believe their excuse is true. However, in either cause you are likely to give the impression that you accept your friend's excuse and move on. This action helps to maintain harmony, just as your friend was helping to maintain harmony by offering an explanation.

Having more interpersonal knowledge often makes relationships harder to manage as Simmel (1950) put it –"relationships … presuppose a certain ignorance and a measure of mutual concealment." It is not always desirable for all parties to be able to account accurately and precisely, therefore leading to participants deliberately making their actions ambiguous.

In IM there is the finding that recipients report a feeling of plausible deniability (Nardi, Whittaker, & Bradner, 2000). Plausible deniability being that recipients rely on the sender's lack of information to excuse a lack of responsiveness. Since the sender doesn't know whether the recipient is there or not, it is not interpreted as rude or unresponsive when no response is given. The person at the other end has to assume that

the other participant had a reason to stop responding since they have no visibility into the other participant's environment.

Conclusion

The three areas that seem to make instant messaging so popular are areas that adding video would actually take away from. For depersonalization, video will obviously make the person personalized. There would no longer be anonymity; the person would become publicly self-aware. If the purpose of using instant messaging was to hide behind the system, to not be aware to others, then adding video would actually destroy that whole purpose, as long as it was two-way.

With ambiguity, plausible deniability would be a lot harder to accomplish, as other people would be able to see whether you are actually there or not. There is less ambiguity which is not always such a good thing.

Lastly for interactivity. Both text chat and video are a visual medium, so putting them together will cause a competition for the users' attention. Once the chat becomes greater than two people, this attention demand will be even higher. An example of where this becomes a problem is with an attempt at video chat conferencing by Scholl, McCarthy, Sasse, and Parnes (2005).

A challenge when delivering multiparty video is screen real estate and bandwidth. With the system developed by Scholl, McCarthy, Sasse, and Parnes (2005) a person automatically appears in the focus window whenever they send a chat message. They reasoned that watching people type would be tedious and would not include much emotional communication since most users would have their attention focused on the keyboard while typing. As a result, they decided to switch the video to a person immediately after they sent a message, since they predicted that emotional expressions were more likely to occur at this time.

Some problems that were noticed were that with large group chat multiple conversational thread overlapped and there was a very rapid turnover of messages, there was also a problem with the visual separation of video and chat with people trying to follow both.

Overall, even though video is a good medium for transmitting social cues and affective information, the way in which it transmits this information goes against what makes instant messaging popular; therefore, video and instant messaging would not be a good combination for social chatting.

There are still possibilities within this area. As mentioned previously about motion providing more information about emotions than still images, even when the motions were of balls and not actual human, leads to the possibility of displaying emotion without necessarily needing to show the actual person. This is just one more idea among many that have been though up for how to display emotion in instant messaging. Ideas have ranged from a haptic device to avatars. At least, it should be a little clearer that going the route of video is not the way to go to keep instant messaging popular and social.

Reference List

Aoki, P.M. & Woodruff, A. (2005). Making space for stories: Ambiguity in the design of personal communication systems. *Proceedings of CHI 2005.*

Arlington, 1996.

Aronson, E. (1992). The return of the repressed: Dissonance theory makes a comeback. *Psychological Inquiry, 3,* 303-311.

Aronson, E. (1998). Dissonance, hypocrisy, and the self-concept. In E. Harmon-Jones & J.S. Mills (Eds.), *Cognitive dissonance theory: Revival with revisions and controversies*. Washington, DC: American Psychological Association.

Aronson, E., Wilson, T.D., Akert, R.M. (2004). *Social Psychology* (4th Ed.). Upper Saddle River, New Jersey: Pearson Education, Inc.

Bassili, J.N. Emotion Recognition. *Journal of Personality and Social Psychology, 37,* (1979), 2049-2058.

Baumeister, R.F. (Ed.). (1993). *Self-esteem: The puzzle of low self-regard*. New York: Plenum.

Brennan, S., & Ohaeri, J. (1999). Why do electronic conversations seem less polite? The costs and benefits of hedging. *Proceedings, International Joint Conference on Work Activities, Coordination, and Collaboration (WACC '99)* (pp 227-235). San Francisco, CA.

Brown, P. & Levinson, S. C. (1978). Politeness: Some universals in language usage. Cambridge, UK: Cambridge University Press.

Bruce, V. The role of the face in communication. *Interacting with Computers, 8,* (1996), 166-176.

Buss, A.H., Self-consciousness and social anxiety. Freeman,San Francisco CA, 1980.

Carver, C.S. and Scheier, M.F., Attention and self-regulation: A control-theory approach to human behavior. Springer- Verlag, New York NY, 1981.

Cherny, L., *The MUD Register: Conversational Modes of Action in a Text-Based Virtual Reality*. P h . D .

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

Clark, H.H., & Carlson, T. (1982). Hearers and speech acts. *Language, 58(2),* 332-373.

Clark, H.H., & Marshall, C.R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B.L. Webber, & I.A. Sag (Eds.), *Elements of discourse understanding* (pp. 10-63). Cambridge, England: Cambridge University Press.

Clark, H.H., & Schaefer, E.F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes, 2(1),* 19-41.

Clark, H.H., & Schaefer, E.F. (1989). Contributing to discourse. *Cognitive Science, 13*, 259-294.

Clark, H.H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22,* 1-39.

Colston, H.L. & Schiano, D.J. Looking and Lingering as Conversational Cues in Video-Mediated Communication. Proceedings ofCH11995 (Denver, CO, May 1995)

Communication/La revue électronique de communication 6(3), 1996b.

Computer Mediated Communication 1(2), 1995.

Condon, S.L. and C.G. Cech, "Discourse Management Strategies in Face-To-Face and Computer-Mediated Decision Making Interactions", *Electronic Journal of*

Connell, J.B., Mendelsohn, G.A., Robins, R.W., & Canny, J. (2001). Effects of communication medium on interpersonal perceptions: Don't hang up on the telephone yet!.

Daft, R. L. and R. H. Lengel, "Information Richness: A New Approach to Managerial Behavior and Organization

Daft, R. L. and R. H. Lengel, "Organizational Information Requirements, Media Richness, and Structural

Danet, B., L. Ruedenberg-Wright and Y. Rosenbaum- Tamari, "Smoking Dope at a Virtual Party: Writing, Play and

Davis, M.H. and Franzoi, S.L., Self-awareness and selfconscioussness, in Personality, Derlega, Winstead, and Jones, (eds.), 1991.

Design", *Research in Organizational Behavior*, volume 6, B.M. Staw and L.L. Cummings (Eds.), JAI Press, Greenwich, CT, 1984, pp. 191-233.

Determinants", *Management Science* 32, 1986, pp. 554- 571.

dissertation, Stanford University, 1995.

Dubrovsky, V., Kiesler, S., & Siegel, J. (1983, October). *Human factors in computer-mediated communication.* Paper presented at the meeting of the Human Factors Society, Baltimore, MD.

Ehrlich, S.M., Schiano, D.J., & Sheridan, K. (2000). Communicating Facial Affect: It's Not the Realism, It's the Motion. Proceedings of CHI 2000.

Harter, S. (1993). Causes and consequences of low self-esteem in children and adolescents. In R.F. Baumester (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 87-116). New York: Plenum.

Herring, S. C. (1999). Interactional coherence in CMC. *Proceedings of the 32nd Hawaii International Conference on System Sciences.* IEEE Press.

Holtgraves, T. M. (2002). Conversational Structure. In T. M. Holtgraves, *Language as Social Interaction* (pp. 89-120). Mahwah, NJ: Lawrence Erlbaum Associates.

Isaacs, E.A., & Clark, H.H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology, 116,* 26-37.

Jefferson, G. (1972). Side Sequences. In D. Sudnow (Ed.), *Studies in social interaction* (pp. 294-338). New York: Free Press.

Kiesler, S. and Sproull, L., Group decision making and communication technology, Organizational Behavior & Human Decision Processes, 52, (1992), pp. 96-123.

Kiesler, S., Siegel, J., and McGuire, T.W., Social psychological aspects of computer-mediated communication, American Psychologist, 39, (1984), pp. 1123-1134.

Kiesler, S., Zubrow, D., Moses, A., & Geller, V. (1985). Affect in computer-mediated communication: An experiment in synchronous terminal-to-terminal discussion. Human-Computer Interaction, I, 77-104.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108,* 480-498.

Language, H. Coleman (Ed.), Mouton de Gruyter, New York, 1989, pp. 251-266.

Lenhart, A. & Shiu, E. (2004). How American use instant messaging. Retrieved April 28, 2006 from http://www.pewinternet.org/PPF/r/133/report_display.asp

Lewin, K. (1943). Defining the "field at a given time." *Psychological Review, 50,* 292-310.

Lewis, D.K. (1969). *Convention: A Philosophical study.* Cambridge, MA: Harvard University Press.

Liseui, C. & Schiano, D.J. Automatic facial expression interpretation. *Pragmatics,* in press.

Lunsford, W., "Turn-Taking Organization in Internet Relay Chat", Unpublished ms, University of Texas at

Marvin, L-E., "Spoof, Spam, Lurk and Lag: The Aesthetics of Text-Based Virtual Realities", *Journal of*

Matheson, K. and Zanna, M.P., Computer-mediated communications: The focus is on me. Special Issue: Computing: Social and policy issues, Social Science Computer Review, 8, (1990), pp. 1-12.

McLaughlin, M.L., *Conversation: How Talk Is Organized,* Sage Publications, Beverly Hills, 1984.

Murray, D.E., "When the Medium Determines Turns: Turn-taking in Computer Conversation", *Working With*

Nardi, B.A., Whittaker, S. and Bradner, E., "Interaction and Outeraction: Instant Messaging in Action," *Proc. CSCW 2000*, ACM (2000), 79-88.

Performance on Internet Relay Chat", *Network and Netplay: Virtual Groups on the Internet,* S. Rafaeli, F. Sudweeks, and M. McLaughlin (Eds.), AAAI/MIT Press, Cambridge, MA,1997.

Pyszczynski, T.A., Solomon, S., Greenberg, J., & Stewart-Fouts, M. (1995). The liberating and constraining aspects of self: Why the freed bird finds a new cage. In A. Oosterwegel & R.A. Wicklund (Eds.), *The self in European and North American culture: Development and processes* (pp. 357-373). Dordrecht, Netherlands: Kluwer.

Reid, E.M., *Electropolis: Communication and Community on Internet Relay Chat.* Senior Honours thesis, University of Melbourne, Australia, 1991.

Rheingold, H., *The Virtual Community: Homesteading on the Electronic Frontier*, Addison-Wesley, Reading, MA, 1993.

Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language, 50,* 696-735.

Schegloff, E.A. (1972). Notes on a conversational practice: Formulating place. In D. Sudnow (Ed.), *Studies in social interaction (*pp. 75-119). New York: Free Press.

Schegloff, E.A., & Sacks, H. (1973). Opening up closings. *Semiotica, 8,* 289-327.

Schelling, T.C. (1960). *The strategy of conflict.* Oxford: Oxford University Press.

Siegel, J., Dubrovsky, V., Kiesler, S., & McGuire, T. (1983). *Group processes in computer-mediated communications*. Manuscript submitted for publication.

Simmel, G., "The Secret and the Secret Society," in *The Sociology of Georg Simmel*, Free Press, New York, 1950, 305-376.

Stone, J. (1998). The self concept in dissonance theory. In E. Harmon-Jones & J.S. Mills (Eds.) *Cognitive dissonance theory: Revival with revisions and controversies*. Washington, DC: American Psychological Association.

Thibodeau, R., & Aronson, E. (1992). Taking a closer look: Reasserting the role of the self-concept in dissonance theory. *Personality and Social Psychology Bulletin, 18,* 591-602.

Tice, D.M. (1993). The social motivations of people with low self-esteem. In R.F. Baumester (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 37-53). New York: plenum.

Walther, J.B. and Burgoon, J.K., Relational communication in computer-mediated interaction, Human Communication Research, 19, (1992), pp. 50-88.

Walther, J.B., "Computer-Mediated Communication: Impersonal, Interpersonal, and Hyperpersonal Interaction", *Communication Research 23(1),* 1996, pp. 3-43.

Werry, C.C., "Linguistic and Interactional Features of Internet Relay Chat", *Computer-Mediated Communication: Linguistic, Social and Cross-Cultural Perspectives*, S. Herring (Ed.), John Benjamins, Amsterdam, 1996, pp. 47- 63.

Whittaker, S., & O'Conaill, B. (1997). The role of vision in face-to-face and mediated communication. In K. Finn, A. Sellen & S. Wilbur (Eds.) *Videomediated communication (*pp. 23-49). Mahwah, NJ: Erlbaum

Whittaker, S., Brennan, S., & Clark, H. (1991). Co-ordinating activity: An analysis of computer supported co-operative work. In *Proceedings of CHI '91 Human Factors in Computing Systems* (pp.361-367). New York: ACM