

## Does Culture Interact with Media Richness? The Effects of Audio vs. Video Conferencing on Chinese and American Dyads

*Leslie D. Setlock, Pablo-Alejandro Quinones, Susan R. Fussell*  
*Human-Computer Interaction Institute*  
*Carnegie Mellon University*  
*5000 Forbes Avenue*  
*Pittsburgh, PA 15213 USA*  
*+1 412-268-7125*  
*[lsetlo, paq, sfussell]@cs.cmu.edu*

### Abstract

*Eastern and Western cultures differ along several dimensions affecting computer-supported collaborative work. We consider one such dimension, low context (requiring little situational information) or high context (requiring substantial situational information) communication style. Specifically, we report on a laboratory study comparing communication and performance of low-context American dyads, high-context Chinese dyads, and mixed American-Chinese dyads on a negotiation task under two possible media conditions: audio conferencing or video conferencing. Although theoretical cultural work and some prior research suggest that high-context Chinese dyads can benefit from the visual cues available in video-enabled systems, we found little support for this hypothesis. There were no effects of culture or medium on conversational efficiency. We did find differences in word usage and quality of interaction between the groups, suggesting potential impact on long term collaborations. We discuss some of the implications of these findings for a theoretical understanding of culture and collaborative work.*

### 1. Introduction

In their seminal paper on why distance matters, Olson and Olson [16] argued that, “possibly the single biggest factor that global teams need to address is culture differences,” and they go on to describe a number of important ways in which cultural differences can impact the success of a collaboration. One important way culture can affect collaboration is by shaping people’s strategies for and expectations about interpersonal communication (e.g., [7]). Some cultures, for example, rely more heavily on nonverbal cues than do others [8]. Problems can arise in cross-

cultural interactions due to mismatches between how people expect their partners to communicate and how they actually do communicate.

Furthermore, cultural differences in communication styles may be differentially influenced by the use of computer-mediated communication (CMC) tools that provide or fail to provide key cues like tone of voice or facial expressions (e.g., [20]). To make good choices about media to use in multinational teams, we need an indepth understanding about how features of communication media interact with culture to influence interpersonal communication. In the current study, we examine two features—the availability of auditory and visual cues—on interactions in American dyads, Chinese dyads, and cross-cultural American-Chinese dyads.

In the remainder of this paper, we first describe the theoretical framework for our research. Then in Section 2, we provide an overview of a laboratory study we conducted to examine interactions between culture and media on communication efficiency and content, participants’ assessments of the quality of their interaction, persuasion, and task performance. In Section 3, we describe our experimental methods, and in Section 4 we present our results. We conclude in Section 5 with a discussion of the implications of our findings and directions for future work.

#### 1.1 Culture and communication

Work in cross-cultural psychology, anthropology, and related fields has demonstrated several differences between cultures that might influence collaborative work. These potential differences include power structures, collectivism vs. individualism, social mobility and many others. One important dimension is

whether a culture is *high context* or *low context* in terms of the amount of contextual information needed for effective communication. According to Hall [8], low-context, typically Western, cultures communicate primarily through verbal channels. High-context, typically Eastern, cultures rely heavily on situational information (e.g., nonverbal behavior, interpersonal relationships).

Whether a culture tends to be low vs. high context has been shown to affect how communicators *ground* their utterances, or ensure that they have been understood as intended [4]. Speakers and listeners work together through processes such as clarification, elaboration, and repair of errors to successfully ground their utterances. Each of these grounding processes can be done by direct verbal means (e.g., the listener can ask a question), or less directly via nonverbal behaviors (e.g., the listener can give a quizzical look). One might expect on a theoretical basis that grounding among members of high-context cultures would rely more heavily on indirect, nonverbal strategies. In a study of verbal grounding strategies in Canadian, Chinese, and mixed Canadian-Chinese dyads, Li [11] found that the effort put into grounding predicted successful information transmission, but only when a Canadian was the listener. In Chinese dyads, grounding appeared to be a relationship-building strategy rather than a mechanism to enhance information transmission.

## 1.2 Culture and media richness

Media richness refers to the extent to which a communication tool provides contextual support for interaction by allowing for immediate feedback, body language, vocal cues, and the like [5]. In the current study, we focus on three types of information that media can provide: visual cues of others' facial expressions, movements and other nonverbal reactions, vocal cues (tone of voice), and interactivity (immediacy of feedback) [3]; see Table 1. Richer media such as video conferencing have more of these features than leaner media such as Instant Messaging (IM). For members of low-context, Western cultures, studies have found that media richness affects such things as the efficiency of grounding or the formality of language, but rarely influences outcome measures such as task performance (see [25] for a review).

Richer computer-mediated communication tools, such as video conferencing systems, provide added social context by making available information about participants' surroundings, attention, gesture and demeanor (e.g., [5][18]). Because high-context cultures theoretically place more value on visual and vocal cues than do low-context cultures [8], researchers have hypothesized that communication among members of

high-context cultures may be more disrupted by the use of leaner media than communication among members of low-context cultures.

**Table 1. Types of contextual information provided by four media conditions.**

Medium	Type of Information		
	Interactivity	Visual Cues	Vocal cues
IM	Yes	No	No
Audio-only	Yes	No	Yes
Video + audio	Yes	Partial	Yes
Face-to-face	Yes	Yes	Yes

Research by Veinott et al. [24] indirectly supports this view. They compared pairs performing a map task via audio vs. video conferencing. Pairs were either American or of mixed cultural background. The results showed that non-native speakers benefited from video whereas native English speakers did not. Veinott et al. suggest that the richer cues to mutual understanding provided by video conferencing (e.g., quizzical looks, halting action, raised eyebrows) were especially helpful to non-native speakers, many of whom were members of Asian cultures. However, because non-native English speakers were always matched with someone from a different culture than their own (i.e., all were in cross-cultural dyads), it is impossible to tell whether the results stem from the specific culture of the participants or from the cross-cultural setting.

Setlock and colleagues [20] compared low-context American dyads performing a negotiation task to high-context Chinese dyads and mixed American and Chinese dyads. Pairs performed two tasks, one face-to-face (a very rich medium) and one via IM (a very lean medium, in which there are no visual or vocal cues). Setlock and colleagues found no effects of media richness on the number of messages required for the task in American dyads. However, there was a significant culture by medium interaction such that Chinese dyads talked longer in the face-to-face condition. They concluded that in a richer communication setting, the Chinese dyads spent time on politeness and other relationship-building activities that the American dyads ignored.

Because Setlock et al. compared a medium with both visual and vocal cues (face-to-face) to a medium with neither of these cues (IM), it is impossible to discern from their results which, if either, of these types of cues is responsible for the medium by culture interaction that they report. Based on Hall's theory,

one would expect members of high-context cultures to be negatively affected by the elimination of both visual cues and vocal cues independently. Visual cues provide important information about social relationships, and vocal cues provide important information about indirect meaning of messages.

Other research suggests that it may be the vocal cues that are primary. For example, D. Zhang and colleagues [26] found that majority influence in a four-person discussion was increased when groups interacted face-to-face compared to when they used either a text-based group support system (GSS), but there was no effect of whether they used the GSS system in the same room (and hence could see one another) or used it remotely (and hence had only textual cues). In addition, they found a significant culture by medium interaction, such that removing vocal cues had a larger impact on Chinese groups. However, people may use visual cues differently when they are also using a text-based system (that is, when these cues do not directly correspond to speaking or listening to messages), we hypothesized that for members of high-context cultures, communication would be more disrupted when both vocal and visual cues were removed (audio conferencing) than when using video conferencing.

## 2. Overview of the current study

Based on the research presented above, we hypothesized that visual cues would be more important for high-context cultures than for low-context cultures. To test this hypothesis, we compared communication and performance on a negotiation task using audio- and video-conferencing for American dyads, native Chinese dyads, and mixed American-Chinese dyads.

We collected five sets of dependent measures: (a) *conversational efficiency*, as indicated by total words and total speaking turns per task; (b) *conversational content*, including word category frequencies and more detailed analysis of argument structure; (c) *quality of interaction*, as indicated by three factors derived from our post-task and post-experiment surveys—coordination, satisfaction with their performance, and frustration; (d) *persuasion*, as indicated by the absolute agreement between pairs' individual post-discussion rankings; and (e) *task performance*, as measured by the agreement between pairs' joint rankings and expert rankings. We anticipated no effect of medium for low-context American pairs but that the high-context Chinese pairs would benefit from the visual cues available in the video connection. Cross-cultural pairs were expected to be intermediate.

## 2.1 Conversational efficiency

Based on the discussion above, we anticipated that whether a pair was conversing via audio or video conferencing would have no impact on American dyads, but audio-only would hinder grounding in both Chinese dyads and mixed American-Chinese pairs. We also anticipated that there would be a main effect of medium, such that all pairs would require more words to complete the task in the audio condition. Finally, we expected an interaction between culture and medium, such that Chinese pairs were more negatively affected by the removal of visual cues in the audio conferencing condition than were American pairs:

*1a. AA pairings will complete the task with the fewest words in both the audio and video conditions.*

*1b. AC pairings will have the most difficulty establishing common ground and thus require the most words in both conditions.*

*1c. All pairs will require more words to complete the task in the audio condition than in the video condition.*

*1d. There will be an interaction between media condition and culture group, such that the effects of media on conversational efficiency will be greatest for the CC group and least for the AA group*

## 2.2 Conversational content

In addition to contextualization and group identification issues, cultures differ in other aspects of communicative style. Here, we focus specifically on inclusive language, which would reflect how collaborators view themselves and each other, and language relating to the negotiation process. We focus here on four categories of words: pronouns (subdivided into “I”, “we”, and “you” pronouns), affective language (e.g., *happy*, *angry*), cognitive mechanisms (e.g., *understand*, *accept*), and social interaction (e.g., *family*, *chat*). In their earlier study, Setlock et al. [20] found use of these word categories to vary by culture and communication medium.

Based on the results established in Setlock et al., we hypothesize:

*2a. Chinese pairs will make greater use of “we” and related group-oriented pronouns whereas American pairs will make greater use of “I” and related singular/individual pronouns*

*2b. Chinese pairs will use the most social language and Americans will use the most cognitive language*

*2c. American and Chinese pairs will engage in greater pronoun use in the video conferencing media.*

## 2.3 Quality of interaction.

The degree to which a medium disrupts fluid communication may be associated with how participants perceive the quality of their interaction. For example, Massey et al. [14] found lower satisfaction levels among Asian participants using an asynchronous text system, which they attributed to the unavailability of prompt feedback and social cues. Similarly, Reinig and Mejias [18] found that GSS-facilitated negotiation resulted in decreased satisfaction with the process for Asian participants as compared to US participants.

Given this combination of factors contributing to satisfaction with both the media and the collaboration, we hypothesize that:

*3a. Chinese participants in AC pairs will have the lowest scores for Quality of Interaction*

*3b. American participants in AA pairs will have the highest scores for Quality of Interaction*

*3c. Scores for Quality of Interaction will be higher for the video mediated condition than for the audio condition.*

## 2.4 Persuasion

Persuasion in the context of a decision-making task refers to the extent to which one team member can convince the other(s) that his/her viewpoint is correct. A sizeable literature on persuasion processes has shown that people are more persuaded by similar others (e.g., [12]). Thus, we can expect that members of same-culture teams will be more open toward their partners' suggestions and ideas than those in cross-cultural teams.

Samarah et al [19] found a higher degree of agreement for Americans than Asians, however that score was due to higher initial disagreement resulting in a greater change. Setlock et al [20], however, found greater levels of persuasion in homogeneously Chinese pairings, which is consistent with D. Zhang et al.'s [26] findings that Chinese were more willing to give into group judgments than were Americans.

Whether or not culture and medium interact to influence persuasion is as yet an open question. Some studies (e.g., [18]) have found main effects of both culture and medium on persuasion but no interaction between the two. In contrast, Anderson and Hiltz [1] found no main effects of group type (culturally homogeneous vs. heterogeneous) or medium but a borderline significant interaction between culture and medium, such that face-to-face heterogeneous groups showed the most consensus and homogeneous

asynchronous groups showed the least consensus. Likewise, D. Zhang et al. [26] found an interaction between culture and persuasion such that Chinese pairs were less likely to give in to group judgments when vocal cues were absent.

We therefore hypothesize that:

*4a. Chinese pairs will demonstrate the highest degree of persuasion.*

*4b. Cross cultural pairs will demonstrate the lowest degree of persuasion*

*4c. Persuasion will be highest in the video-mediated condition.*

## 2.5 Task performance

The interactions between culture and medium that we have described thus far may be expected to affect how well pairs can come up with the correct solution in our negotiation task. When pairs cannot communicate well, and can't successfully persuade one another, they may be less likely to come upon the correct solution. Thus, we expect that:

*5a. Participants will perform better in the video conferencing condition than in the audio conferencing condition.*

*5b. Chinese pairs' performance will be more enhanced by video than American pairs..*

## 3. Method

### 3.1 Design

Pairs of participants negotiated a modified version of two decision-making tasks (the Desert Survival Task and the Arctic Survival Task), one via video conferencing and the other via audio conferencing. Pairs were comprised of two native-born American participants (AA), two native-born Chinese participants (CC), or one American and one Chinese participant (AC). The order of tasks and media conditions was counterbalanced.

The experiment was conducted in English, even for the Chinese pairs, so that we could compare conversational efficiency and content across all three cultural groups (AA, AC, and CC). However it should be noted that in a previous study in which we replicated Setlock et al.'s [20] comparison of dialogues in IM vs. face-to-face using Chinese dyads speaking in Chinese, we found no differences between these results and the original Setlock et al.'s findings in which all participants had conversed in English [21].

### 3.2 Participants

Thirty-four pairs of participants (10 AA, 12 AC, 12 CC) were recruited from two local universities. Half of the participants were citizens of the United States who spoke English as their first language. The remaining participants were visiting students from the People's Republic of China whose first language was Chinese, and who had been in the United States for fewer than five years. The 5-year criteria was established in order to select participants who would have demonstrated conversational levels of English proficiency through regular interaction, but who would still maintain active involvement with Chinese culture and have strong ties to Chinese community.

Chinese students were recruited through two local chapters of the Chinese Student and Scholars Group, as well as other Chinese cultural groups. The Chinese students were all fluent, or nearly fluent, in spoken and written English.

### 3.3 Materials

Participants negotiated a shortened version of the Desert and Arctic survival tasks by Human Synergistics [9]. The goal of these tasks is to rank a set of salvaged items in order of importance for the survival of the team. Participants were given a scenario (e.g., they were on a plane that went down in the arctic) and a list of six salvaged items. They first ranked the items individually, then negotiated and completed a joint ranking with their partner, and finally completed a second individual ranking. Items included a "gallon can of maple syrup" and a "hand ax" (Arctic Survival Task) and "Book 'Edible Plants of the Desert'" and "loaded .38 caliber pistol" (Desert Survival Task.) Items were ranked from 1 (most important) to 6 (least important.)

Participants also completed a post-task questionnaire following each task, which assessed their experience with the communication media and partner. The questionnaire included such items as "This method of working together was effective," and "We disagreed often." Items were rated from 1 (strongly disagree) to 7 (strongly agree).

### 3.4 Equipment

Web cameras were located above each participant's monitor, which sent a head and shoulders view of that participant to his/her partner during the video condition. Video was displayed full-screen on a color 13-inch monitor located directly in front of each participant (see Figure 1). Sony wireless microphones

were used to record audio in both conditions. Because the participants were actually in close proximity, though visually separated by a barrier, audio technology for transmission was not required



**Figure 1. Arrangement of equipment in the video conferencing condition. Cameras were oriented to allow eye contact.**

### 3.5 Procedure

Participants were brought to the laboratory and instructed about their task and the media (microphone and, web cameras) used in each condition. Following the instructions, they were situated at separate desks in the same room, separated by a divider. For each trial, participants first read the scenario and completed their individual item ranking of the salvaged items. They then had 20 minutes in which to discuss and complete the joint ranking (a single, mutually agreed-upon ranking), after which they completed the final individual ranking. Each trial was followed by a post-task questionnaire. After both trials were completed, participants completed a post-experimental questionnaire. They were then debriefed, compensated and dismissed.

## 4. Dependent measures

### 4.1 Conversational efficiency

The sessions were transcribed. We then used Adam Kramer's Text Analysis and Word Count (TAWC) program sources [10] to automatically count the number of speaking turns per participant per trial and the total number of words spoken.

## 4.2 Conversational content

We also used the TAWC program to calculate the percentage of words in several content categories shown by Pennebaker and his colleagues (e.g., [17]) to reveal important dimensions of interaction. We anticipate that rates of use for these words will be identified with different expectations of the interactions – i.e., whether the participants see themselves as a team or as two individuals, as involved in either a task- or relationship-oriented communication, and with what degree of emotional/social involvement with each other.

## 4.3 Quality of interaction

In order to gauge their perception of the interaction, participants responded to questions about the success of their collaboration on the post-task questionnaire using a scale from 1 (strongly disagree) to 7 (strongly agree). Factor Analysis with Varimax rotation indicated one four-item factor corresponding to how well the pair collaborated (“my partner was responsive to my ideas,” “my partner treated me fairly”). These questions formed a reliable scale ( $\alpha = .83$ ) and were averaged to create one variable reflecting the participants’ perceptions of the success of the collaboration.

## 4.4 Persuasion

We assume that persuasion is reflected by changes in the size of the differences between participants’ pre- and post-discussion rankings. Initial agreement was computed by summing the absolute values of the differences in scores for each item prior to the pairs’ discussions. Final agreement was computed by summing the absolute values of the differences in scores for each item after the pairs’ discussion. Scores for agreement range from 0 (perfect agreement) to 18 (perfect disagreement).

## 4.5 Task performance

Task performance is a feature of the objective “correctness” of a collaborative effort. In this case, the goal of the task was to rank the items in order of most to least important. The correctness of their joint answers was determined by comparing their scores to rankings provided by survival experts.

Performance was measured by the distance between the pairs’ joint rankings and expert rankings provided by Human Synergetics. Performance was computed as the sum of the absolute value of the difference in

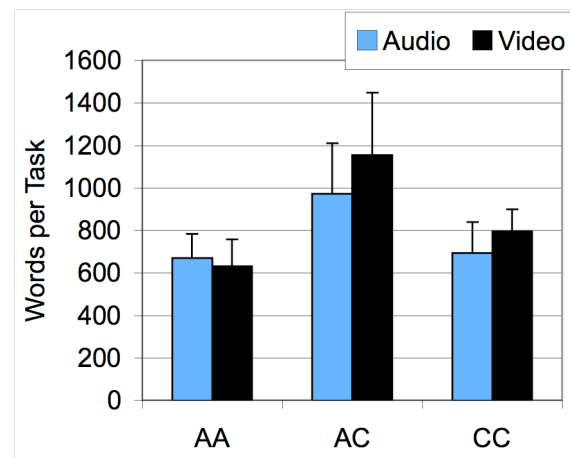
scores for each item. Similarly to the persuasion rankings, 0 on this measure indicates perfect agreement with Human Synergetics’ expert rankings whereas 18 indicates perfect disagreement with the expert rankings.

## 5. Results

Results were analyzed using Mixed Models Analyses of Variance in which pair was a random factor and medium (audio vs. video conferencing) was a fixed factor. Preliminary analyses indicated no effect of trial or task, so we excluded these variables from the final analyses reported here.

### 5.1 Conversational efficiency

Contrary to Hypotheses 1a-1d, we found no main effects of culture or medium, nor an interaction between the two, on the total number of words required for the two tasks. Pairs used a mean of 784 words (SD = 549) in the audio condition and a mean of 873 (SD = 646) in the video condition, across cultures. Means by culture group and media condition are shown in Figure 2, and demonstrate that, overall, the cross-cultural pairs used the greatest number of words, followed by the homogenous Chinese pairs.



**Figure 2. Mean number of words per pair by media condition. Error bars represent the standard errors of the means.**

Although the large standard deviations quite possibly limited our ability to find significance, they are interesting in that they show considerable variability between pairs in how much they discussed their tasks. This was particularly marked for the cross-cultural pairings. There were also no main effects of culture or

media or interactions for the amount of time it took to complete the task.

### 5.2 Conversational content

Main effects of culture were found on the proportions of several message content variables, supporting Hypotheses 2a and 2b. First, we examined the use of different categories of pronouns (“I”, “we” and “you” pronouns) as a percentage of the total number of pronouns used. There was a significant effect of culture on “we” pronouns ( $F [1, 25] = 3.9, p < .05$ ) and a borderline effect of culture on “you” pronouns ( $F [1, 24] = 2.68, p = .09$ ). As shown at the top of Table 2, “we” pronouns were more common in pairs that included at least one Chinese member, whereas “you” pronouns were most common in the American pairs.

Consistent with Hypothesis 2b, there was a borderline effect of culture on the use of terms reflecting cognitive mechanisms ( $F [1, 24] = 3.32, p = .05$ ) showing greater use of words relating to thinking through the process, such as “understand” and “accept,” by American pairs.

**Table 2. Mean “I”, “we”, and “you” pronouns as a percent of pronouns (top) and mean affective, cognitive and social words as a percent of words (bottom) by culture group (standard deviations in parentheses).**

	AA	AC	CC
Pronouns:			
I	39.4 (10.5)	42.4 (9.8)	39.5 (10.2)
We	6.5 (7.6)	13.4 (8.2)	14.8 (7.9)
You	30.1 (9.5)	22.1 (5.7)	26.6 (9.7)
Affective	3.7 (0.9)	3.3 (0.7)	3.7 (1.3)
Cognitive	10.3 (1.6)	9.2 (1.9)	8.6 (1.2)
Social	6.4 (1.3)	6.3 (1.5)	6.7 (1.6)

Contrary to Hypothesis 2c, there were few effects of communication medium on the use of word categories. There was a borderline significant tendency for participants in the video condition to use more social terms ( $F [1, 25] = 3.45, p = .08$ ). There were no interactions between culture and medium.

### 5.3 Quality of interaction

Each participant rated the quality of their collaboration with their partner after each of their tasks. Scores ranged from 1, reflecting poor collaboration, to 7, reflecting excellent collaboration. Scores were analyzed using a Mixed Models ANOVA

in which participant within pair was a random variable and culture group and communication medium were fixed effects. Results showed that the culture groups differed in their perceptions of the effectiveness of the collaborations ( $F [2, 65] = 3.62, p < .05$ ), with Chinese participants feeling less effective in their task. Contrary to Hypothesis 3a and partially consistent with Hypothesis 3b, post-hoc t-tests showed that the Chinese dyads gave their interactions somewhat poorer rankings than did the American dyads ( $M_s = 5.74$  vs.  $6.34, t = -1.80, p = .08$ ).

### 5.4 Persuasion

Persuasion was determined by the absolute difference in ranks in participants’ post-discussion individual rankings such that a score of 0 indicated perfect agreement. This measure describes the extent to which the negotiation can be seen as having affected the participants’ perceived item importance rank. Contrary to Hypotheses 4a-4c., we found no main effects of culture or medium and no interaction between the two on the degree to which participants came to agree on the ranking of the items. Mean post-discussion agreement was 2.00 for the AA pairs ( $SD = 2.51$ ), 1.75 for AC pairs ( $SD = 1.98$ ) and 1.92 for CC pairs ( $SD = 2.24$ ). Thus, unlike the American pairs in Setlock’s study, our American pairs came to quite strong agreement. Final agreement between pairs was slightly but nonsignificantly higher in the video condition ( $M = 1.59, SD = 2.13$ ) than the audio condition ( $M = 2.18, SD = 2.28$ ).

### 5.5 Task performance

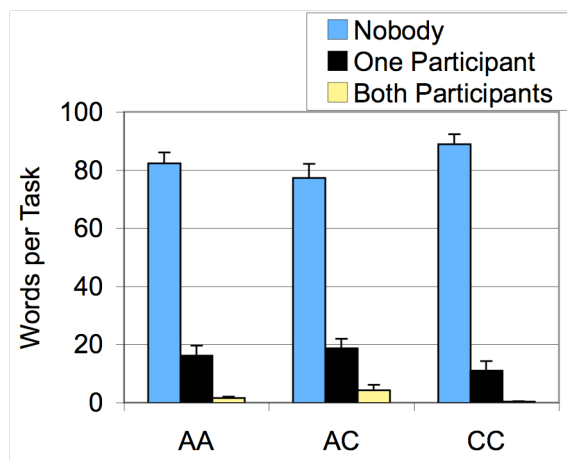
Task performance was assessed by comparing pairs’ joint rankings to the rankings provided by Synergistic’s experts. Like Setlock et al.’s pairs, our pairs were quite far off from experts in their final rankings ( $M = 9.49, SD = 3.50$ ). It appears that the results of different styles of argumentation and negotiation are present in how participants feel about the collaboration and how much they agree, but not on the correctness of their choices according to an outside judge. Contrary to Hypotheses 5a and 5b, there were again no main effects of culture or medium and no interaction between the two.

### 5.6 Gaze behavior

Because we found very few effects of medium on participants’ communication or performance, we went back through the video recordings for each session and coded subjects’ gaze at the video monitors. We coded

when no one was looking at the monitor, one participant was looking at the monitor, and both participants were simultaneously looking at the monitor (thereby establishing mutual gaze). As a control for participants' general looking behavior, we coded gaze for both the audio and video sessions. Scores were converted into percentages of gaze in each category (no one looking, one person looking, both looking).

In the audio condition, no one was looking at the monitor for 96% of the time. In the video condition, as can be seen in Figure 3, much of the time nobody was looking at the monitor and very little of the time did participants establish mutual gaze. ANOVA confirmed that the differences between culture groups were not significant ( $F [2, 32] = 1.37, p > .10$ ). It thus appears that participants rarely took advantage of the visual cues provided by the video conferencing system and that members of high context CC pairs were no more likely to use these cues than members of AA or AC pairs.



**Figure 3. Percentage of time in the Video Condition in which nobody was looking at the monitor, one participant was looking at the monitor, or both participants were looking at the monitor. Error bars represent the standard errors of the means.**

## 6. Discussion

This study investigated the effect of two features of rich media—tone of voice and visual cues—on the communication and performance of low-context, high-context, and cross-cultural dyads. Our general hypothesis was that high-context Chinese dyads would benefit from the added visual cues in video conferencing over audio conferencing, whereas low-context American dyads would not. Instead, we found

few cultural differences between conversations in audio versus video conferencing systems.

A theoretical implication of our findings is that the relationship between culture and media richness is not linear. When comparing a very lean medium (IM) to very rich medium (face-to-face conversation), Setlock et al. [20] found an interaction between culture and medium, and media effects on Chinese dyads were later replicated using Chinese participants speaking in Chinese [21]. In the current comparison between two media of intermediate richness, however, we found no such interaction.

One likely explanation for our results is that it is the auditory features of face-to-face interaction that are most important for high-context cultures, at least for the type of task we used here. This is consistent with D. Zhang et al.'s [26] findings with respect to culture and media richness in their study of GSS systems.

Another possibility is that, for Chinese participants, written language varies more dramatically from spoken language than for American participants. This would explain the lack of a linear effect of increased richness corresponding to increased effectiveness. Because the same effects of typing in IM were found when Chinese participants were typing in English [20] and when they were typing in Chinese [21], it is clearly not simply the effort involved in typing English that affects the way the high-context Chinese communicate.

### 6.1 Culture vs. native language

A puzzling result is the failure to replicate Veinott et al.'s [24] findings that video benefited non-native English speakers but not native speakers. We see two possible explanations for this lack of replication that we will attempt to disentangle in future research. One is that our Chinese participants were more fluent in English than Veinott et al.'s, at least within the task domains of the studies. Alternatively, the results may stem from differences between the participant combinations used in the two studies. In Veinott's study, non-native English speaking participants were paired with members of different non-native English speaking cultures. Each non-native speaker was therefore also in a culturally heterogeneous pairing, confounding native language use and cultural diversity. In our study, in contrast, Chinese participants were always paired with either another Chinese participant or an American participant. Thus, it is possible that non-native language speakers in culturally heterogeneous conversations might benefit from the added contextual information provided by video conferencing over audio conferencing.

The task and environment's suitability to video-conferencing may also play a role. Since neither



participant in our experiment looked up at the video often, the odds of them looking simultaneously, thereby catching each others' eye for any length of mutual gaze time, was unlikely. It is possible that reorganizing the camera and desk angles or putting the experimental materials on a monitor adjacent to the camera might encourage gaze and result in greater use and subsequent benefits of the visual information.

## 6.2 Collaboration, communication and relationships

Curiously, the American dyads in our study spoke much longer than American dyads in the face-to-face condition of Setlock et al.'s study [20]. In addition, the highest overall rate of word usage was by the cross-cultural pairings, unlike in Setlock et al.'s study where the CC pairs used the greatest number of turns. A preliminary examination of the dialogues showed that American dyads also discussed the task in greater depth than the cursory analysis Setlock et al. report for American dyads in their study. Interestingly, we found that Americans benefited in terms of persuasion from this deeper discussion of the task in much the same way as the Chinese pairs did in Setlock et al.'s study. This may suggest that, while a longer, more detailed communication style is more natural for Asian collaborators, the benefit of increased agreement can be achieved by both cultures if such discussion is facilitated.

We did find main effects in the level of pronoun use by culture that are consistent with the cultural dimension of individualism-collectivism (e.g., [23]). Those pairs including a Chinese participation (i.e., the AC or CC pairs) engaged in significantly greater use of "we," and the homogeneous American pairs used the most "you" pronouns. There was no interaction between pronoun use and media, suggesting that feelings of "we"-ness are not dependent on visual awareness of one's partner. Coupled with the trend for increased social language in the video condition, this suggests richer media may better support cross-cultural collaboration by encouraging a greater sense of social relationship.

Consistent with Reinig and Mejias' [18] work, we found that Chinese pairs experienced lower levels of satisfaction when communicating without visual awareness. Although this did not significantly impact upon their task completion or efficiency, it is worth considering the long-term collaboration implications where the media is significantly more socially fulfilling for one group than another.

## 6.3 Conclusions and further directions

Although many of our original hypotheses were not supported by our results, we believe our findings raise several interesting questions for future research. First, the question of whether media richness follows the same pattern, and in the same degree, for different cultures would have bearing on the choice of communication media. Our study, in light of prior text vs. face-to-face research, suggests that written media serves a very different role for Chinese than for Americans. While media richness maintains an approximately linear relationship for American participants, for Chinese participants the biggest difference exists between text-based and vocal, with minimal benefit of increased context among different vocal media.

In addition, it would be worth conducting future work on media of varying richness with diverse cultures using different tasks, to see if there is a threshold at which point the social context cues become more important. Although we did not find efficiency and performance differences in our study, there are other important dimensions of interpersonal interaction that may show interactions between culture and medium. For example, Q. Zhang [27] has found differences between media in terms of how well they support trust development in American versus Chinese dyads. Our task could be completed with more or less personal investment. In a task requiring more social involvement, participants may need to rely more on the social cues afforded by media (even if the media does not support them in exactly the way afforded by face-to-face contact.) We plan to expand our future work to include trust and other measures of social relationship quality.

## Acknowledgments

This material is based upon work supported by the National Science Foundation under Grants #0325087 and #0329077. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We thank Sara Kiesler, Susan Straus, Suzanne Weisband, Shalom Schwarz and Romie Littrell, for their helpful input into the design and analysis of this study. We thank Adam Kramer for adapting his TAWC software for our project and for running analyses, and Kristina Izumi, Gail Kusbit, Julie Ludwick, Keiko Wada and Shuang-Qi Zheng for their assistance in running experiments and coding data.

## References

- [1]Anderson, W. N. & Hiltz, R. S. (2001). Culturally Heterogeneous vs. Culturally Homogeneous Groups in Distributed Group Support Systems: Effects on Group Process and Consensus. *Proceedings of the 34th Hawaii International Conference on System Sciences*.
- [2]Brock, T.C. (1965). Communicator-recipient similarity and decision change. *Journal of Personality and Social Psychology, 1*, 650-654.
- [3]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 Press.
- [4]Clark, H. H. & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition, 22*, 1-39.
- [5]Daft, R.L. & Lengel, R.H. (1984). Information richness: a new approach to managerial behavior and organizational design. In L. L. Cummings & B.M. Staw, (Eds.), *Research in organizational behavior 6*, (pp. 191-233). Homewood, IL: JAI Press
- [6]Fussell, S.R., Kraut, R.E., & Siegel, J. (2000). Coordination of communication: Effects of shared visual context on collaborative work. *Proceedings of CSCW 2000* (pp. 21-30). NY: ACM Press.
- [7]Gudykunst, W. & Ting-Toomey, S. (1988). *Culture and Interpersonal Communication*. Newbury Park, CA Sage.
- [8]Hall, E. 1976/1981. *Beyond Culture*. New York, NY: Doubleday/Anchor Books.
- [9]Human Synergistics. [www.humansyn.com](http://www.humansyn.com).
- [10]Kramer, A.D.I, Fussell, S., & Setlock, L. (2004). Text analysis as a tool for analyzing conversation in online support groups. In *Proceedings of CHI 2004*.
- [11]Li, H. Z. (1999a ). Grounding and information communication in intercultural and intracultural dyadic discourse. *Discourse Processes, 28*, 195-215.
- [12]Mackie, D.M. (1986) Social identification effects in group polarization. *Journal of Personality and Social Psychology, 50*, 730-728.
- [13]Matheson, K. (1991) Social cues in computer-mediated negotiations: Gender makes a difference. *Computers in Human Behavior, 7*, 137-145.
- [14]Massey, A., Montoya-Weiss, M., Hung, C. & Ramesh, V. (2001) When culture and style aren't about clothes: Perceptions of task-technology "fit" in global virtual teams. *Proceedings of GROUP 2001* (pp. 207-213). NY: ACM Press.
- [15]Neuwirth, C. M., Chandok, R., Charmet, D., Wojahn, P. & Loel, K (1994). Distributed collaborative writing: A comparison of spoken and written modalities for reviewing and revising documents. *Proceedings of CHI 1994* (pp. 51-57). NY: ACM Press.
- [16]Olson, G. M., & Olson, J. S. (2000). Distance matters. *Human-Computer Interaction, 15*, 139-179.
- [17]Pennebaker, J.W., Francis, M.E., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count: LIWC (2<sup>nd</sup> ed.)*. Mahwah, NJ: Lawrence Erlbaum Associates.
- [18]Reinig, B. A., & Mejias, R. J. (2003). An investigation of the influence of national culture and group support systems on group processes and outcomes. *Proceedings of HICSS 2003*.
- [19]Samarah, I., Paul, S., Mykytyn, P. Jr., & Seetharaman, P. (2003). The collaborative conflict management style and cultural diversity in DGSS supported fuzzy tasks: An experimental investigation. *Proceeding of HICSS 2003*.
- [20]Setlock, L. D., Fussell, S. R., & Neuwirth, C. (2004). Taking it out of context: Collaborating within and across cultures in face-to-face settings and via instant messaging. *Proceedings of CSCW 2004* (pp. 604-613). NY: ACM Press.
- [21]Setlock, L. S., Fussell, S. R., & Shih, Y. Y. (2006, July). *Effects of culture, language and communication medium on conversational grounding*. Annual Meeting of the Society for Text and Discourse, Minneapolis, MN.
- [22]Straus, S., & McGrath, J. (1994). Does the Medium Matter? The Interaction of task type and technology on group performance and member reactions. *Journal of Applied Psychology, 79*, 87-97
- [23]Triandis, H. C. (1989). The self and behavior in different cultural contexts. *Psychological Review, 96*, 506-520.
- [24]Veinott, E., Olson, J., Olson, G. & Fu, X. (1999) Video helps remote work: Speakers who need to negotiate common ground benefit from seeing each other. In *Proceedings of CHI 1999* (pp. 302-309). NY: ACM Press.
- [25]Whittaker, S. (2003). Theories and methods in mediated communication. In A. Graesser,, M. Gernsbacher, and S. Goldman, (Eds.) *The Handbook of Discourse Processes* (pp. 243-286). Mahwah, NJ: Lawrence Erlbaum Associates.
- [26]Zhang, D., Lowry, P., B., Fu, K., Zhou, Li., Adipat, B. (2006). Culture and media effects on group decision making under majority influence. *Proceedings of HICCS 2006*.
- [27]Zhang, Q. P., Sun, X., Chintakovid, T., Ge, Y., Shi, Q., & Zhang, K. (2006). *How culture and media influence personal trust in different tasks*. Paper presented at the HCIC Winter Consortium, Feb. 2006.